认知升级

[美] 理查德·尼斯贝特（Richard E. Nisbett） 著

仲田甜 译



INTELLIGENCE AND HOW TO GET IT

重塑思维与认知，实现自我进化

中信出版集团

版权信息

书名:认知升级

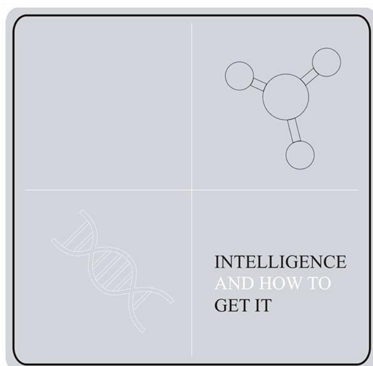
作者:[美]理查德·尼斯贝特

译者:仲田甜

ISBN:9787508679426

中信出版集团制作发行

第1章
测测看，你的智商多少分？



> 心理学家认为智力是一种天生的综合的智能，是靠遗传，而不是通过教育或训练得到的，勤奋或热情也不会对智力产生任何影响。
>
> ——西里尔·伯特爵士及其同事（**1934**年）

小学五年级我们班刚刚开始学习分数时，我有一周没到学校上课，此后我学习算数就开始吃力起来。在接下来的小学生涯里，我的算术成绩再也没有好起来。父母很同情我，安慰我说，我们家的人数学都学得不太好。他们认为数学能力主要是靠遗传得来的，要么生来就有，要么生来就没有。

我的父母恐怕并没有读过有关智力的心理学文献，但他们的观点与之不谋而合。20世纪末，不少专家都认为智力和学习能力在很大程度上受到基因的控制，它们在出生前便被预设好，在正常的环境下或多或少地得以展露。这些专家认为任何为提高智力而做出的努力都不会得到回报。所以，像儿童早期教育这种干预措施并没有取得持久的效果，也就不足为奇了。他们不相信通过改善教育或改变社会就能够使人们变得更加聪明。

不过心理学、遗传学和神经系统科学近期的研究成果，以及目前对教育干预有效性的研究，已经合力推翻了根深蒂固的智力遗传决定论的立场。我

们现在可以肯定的是，智力的环境可塑性非常强。一个没有接受过正规教育的人无论如何也不可能变得非常聪明——无论我们是用智力测试还是别的标准来测量。一个人的智力，及其学业成就和职业成就的高低，很大程度上取决于环境因素，而与基因没有任何关系。

这一新近产生的环境决定论有三条重要的原则：

1．包括学校教育在内的正确的干预措施，可以使人变得更加聪明。当然学校教育还可以被改善得比现在好很多。

2．社会对智力的要求变得更高了，加之文化及教育环境的不断改善，使得所有人口在整体上都变得更加聪明了——但这种聪明与以前的聪明有所区别。

3．经济地位的高低造成了不同阶层的人在智力及学业成就上的差距，但这种差距是可以缩小的。同样地，白人和某些少数族裔之间的此种差距也是可以缩小的。

本书的基本内容很简单，就是关于环境对智力潜能的影响力，更具体地说，是关于学校及文化对于环境的影响作用。越来越多的研究证据，尤其是大量最新的研究证据让我们有理由相信，和几年前的大多数专家相比，我们应该更加乐观地看待个人、群体及社会整体智力提升的可能性。

另一方面，就像有些外行和专家错误地坚持智力主要取决于遗传的观点一样，另一些外行和专家对于能够提升智力和学业表现的方法存在一定误解，有时甚至是过于乐观。本书的目标之一就是要证明究竟何种干预措施才能最有效地提升智力和学业表现。

在接下来的几章里，我会着重阐述不同群体间的社会和文化差异对智力及学业成就的重大影响。社会经济地位较低的人平均智商及成就都不高，这一部分是由环境因素造成的——并且部分环境因素在本质上也是文化因素。而黑人及其他少数族裔的智商和成就不高，则完全是由环境因素造成的。大部分环境因素都与历史上遗留下来的不利条件有关，但有些也和社会习俗有关，而这些社会习俗是可以改变的。

文化也能够促进智力和学习成绩的进步。大体来说，某些文化群体相较于主流的白人群体具有明显的智力优势。这些群体包括带有东亚血统的人和德系犹太人。后面我会讨论这些优势是如何产生的，以及那些想要提高自己智力和学业成就的人是否也能获得其中的一些优势。

最后，我会介绍最新的科学发现所显示的提高智力的方法。

不具备任何专业知识的读者也能很好地理解本书中绝大部分内容，不过熟悉一点儿统计学的知识可能会更有帮助，所以我在附录中对一些统计学术语下了定义。想要加强统计学知识的读者也可以参考这个附录。其中涉及的概念包括：正态分布、标准差、统计显著性、效应量（在标准差术语中）、相关系数、自我选择以及多元回归分析。

多元回归分析，就是对一些自变量进行测量并检验这些自变量与其他因变量的关系。请注意，我对多元回归分析有点儿排斥，因为这样的分析不能正确地显示不同变量之间究竟存在何种程度的因果关系，所以我极少谈到多元回归分析，并一贯持怀疑态度。想要了解我为何对多元回归分析带有偏见的读者可以看看附录一，我在里面做了相关解释。

作为全书的开始，在这一章里我会给出智力的定义，探讨测量智力的方法，提供智力测试可以测量出的两种不同类型的分析型智力的例证，以及探讨智力测试无法测量出的智力类型。同时，我还会检验智商究竟能够在多大程度上预示未来的学业成就和职业成就，智商不能代表的智力类型以及有关动机和性格的一些重要方面。

智力的定义及测量方法

我们可以将琳达·戈特弗雷德森对智力的定义当作一个很好的开始：

> 智力是一种内容非常广泛的脑力活动，其中涉及推理、计划、解决问题、抽象思考、理解复杂概念、迅速学习以及从经验中学习的能力，而不局限于学习书本知识、学术能力或应试技巧。智力反映的是一种能够更加广泛和深刻地理解我们周遭环境的能力——“理解”、“领会”事物，或“决定”要做什么。

智力方面的专家几乎一致认为，智力包括抽象推理、解决问题和获取知识的能力。绝大部分专家也都认为，记忆力和心理速度也是智力的一部分，还有大约一半的专家认为常识和创造力也应算作智力的一部分。

这些定义没有将智力的某些方面包含在内，而这些方面恰恰构成了其他文化中智力的组成部分。发展心理学家罗伯特·斯滕伯格研究了大量不同文化中非专业人士对于智力组成部分的看法。他发现，很多人都将理解他人并产生共鸣等社交特征视为智力的一部分，这在非洲和东亚文化中尤为突出。不仅如此，东亚人对于智力的理解比西方人更强调实际和实用性。西

方人更有可能看重对知识本身的研究，不论这些知识是否有很明显的直接用途。

我们通常都利用智力测试对智力进行测量。顺便提一下，智商（IQ）的“商（Q）”代表的是商数（quotient）。最初的智力测试是为小学生设计的，用智力年龄除以生理年龄得到的就是商数。按照这个定义，如果一个10岁的孩子在测试中的表现相当于一个12岁的孩子，那么这个孩子的智商就是120分；如果一个10岁的孩子在测试中的表现相当于一个典型的8岁孩子，那么其智商就是80分。不过现代的智力测试都是将某个特定年龄群人口的平均智商值定为100分，然后使智商值围绕这个平均值分布，由此获得特定的标准差——通常是15分。因此，如果一个人在智力测试中的表现比他（她）所属年龄群体的平均值高1个标准差，那么他（她）的智商就是115分。

这里举一个例子说明智商相差15分意味着什么。智商为100分的某个人可能会以一般的成绩从高中毕业，然后再上一两年的社区学院。而一个智商为115分的人预计能从大学毕业，进而成为某行业中的专业人士或企业中的高层管理人士。反过来，如果某个人智商为85分，也就是正常值范围内的最低值，他（她）极可能连高中都无法毕业，只能找份技术工人的差事。

虽然设计智力测试是为了预测学习成绩，但我们很快就发现，这种测试所测量的智力与普通人理解的智力有很多地方是重合的。至少，人们对于他人智力水平高低的评定与智力测试得到的结果基本是吻合的。那些被普通人视为智商较高的人，在智力测试中得到的分数也相对较高。

智力测试种类繁多，但是综合性的智力测试之间都没有太大的区别，而任意两种智力测试，即使是测试内容明显不同，其典型相关度也都在0.8~0.9的范围之内。

智力测试有时会测量一些非常具体的技能，例如拼写能力和推理速度。这些高度具体化的测试往往有彼此相关成为集群的倾向。[注] 例如记忆力测试往往都是相互关联的，各种测量视觉和空间感知能力（例如将彩色的积木拼成一个二维的图案）以及测试语言知识的试题（例如词汇）往往也是如此。所有测量你能称之为智力的测试，至少从某种程度上来说都是相互关联的。（就这件事而言，任何被社会认为是好的事物，都和其他好的事物相关联。生活是不公平的。）

表**1–1** 韦克斯勒儿童智力测验量表中使用的小测验

常识：	哪些洲完全位于赤道以南？
词汇：	贬抑是什么意思？
理解：	为什么街道通常都要按照顺序编号？
类同：	树和花具有哪些相似之处？
算术：	如果6个橘子要花费2美元，那9个橘子要花费多少美元？
图画补缺：	指出一幅不完整的图画中缺失的部分
积木图案：	将积木拼成一个两种颜色的图案
图形拼凑：	将拼板摆成一个寻常物件的形状
图片排列：	重新排列一组被打乱顺序的图片，使之能够描述一组有意义的事件
译码：	按照所给的例子将符号和图形一一对应起来

表1–1就是一个智力测试的范例，显示的是韦克斯勒儿童智力测验量表中的各项小测验，用来测试6～16岁孩子的智商。像这样的智力测试，各小测验之间的相关度为0.3～0.6。有一种说法认为，小测验之间具有相关性，就表示有可能存在某种因素，这种因素可以反映一般智力水平的高低，叫作g因素。（“因素”有技术上的含义，但我们没有必要细究。g因素本身与智商得分高度相关，但在某些方面与智商值存在细微的差别，不过这些方面与我们现在的讨论无关。）某些小测验与g因素的相关度要比其他小测验高一些，因此这些小测验的g负荷量较高。例如，词汇测验与g因素的相关度就很高，而译码（按照所给的例子将种种符号对应起来）与g因素的相关度就不高。

智力的两种类型

实际上g因素或一般智力有两个组成部分，其中一个是液态智力，或叫作解决新奇的抽象问题的能力——这种智力需要的是智力活动，而这种智力活动对于人在一生中不断获取的现实生活经验的依赖不大。液态智力通过所谓的“执行功能”的运转得以运用，这些执行功能包括“工作记忆”、“注意力控制”及“抑制控制”。那些为了解决问题而必须经常储存在头脑里的信息，以及那些需要花些力气才能记住的信息，据说都保留在工作记忆中。注意力控制这种能力，不仅要求将注意力集中在有待解决的问题的相关方面，而且能在需要的时候将注意力转移到问题的下一环节。抑制控制是指

抑制具有诱惑性但并不相关的行动的能力。

图1–1中显示的是一道用于测试液态智力的典型问题。这是从瑞文推理测验中摘录出来的。约翰·C. 瑞文在1938年发表了这项测试的第一版。

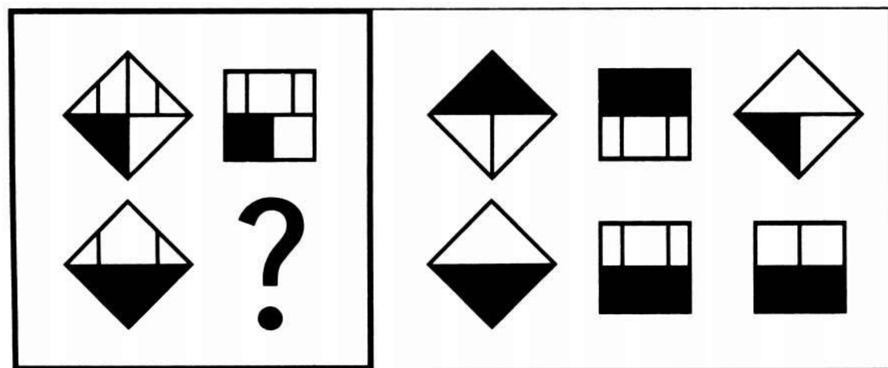


图1–1 类似于瑞文推理测验的一道题

资料来源：Flynn, 2007.

左边方框内上面的两个图案是解决该问题必须要遵循的范例，下面左边的图案指定了要想解决这个问题需要进行哪些变形。右边方框内的6个图案是备选的答案。要解决这个问题，你就要注意到左边方框内左上方的图案是一个菱形，右上方的图案是一个正方形。这就告诉你答案必须是一个正方形。然后你必须要注意到左上方菱形的下半部分被平分成了两部分，左半部分是黑色的；右上方正方形图案的左下半部分也是黑色的。这就告诉你右下方正方形中对应的部分必须和左下方菱形图案对应的部分一致——也就是说，整个下半部分必须都是黑色的。你还应注意到，左上方菱形图案中中间的一条竖线被拿掉了，保留了对称的两条竖线，这就决定了为使图案对称，你必须清除右下方正方形中一条竖线。这样正确的答案就出来了，也就是右边方框内右下角的那个正方形。

在表1–1显示的韦克斯勒儿童智力测验量表的各项小测验中，涉及液态智力较多的是：图画补缺，你需要关注一个图案的所有方面，然后分析其中的哪个部分缺失了；积木图案，你需要处理纯粹抽象的视觉材料；图形拼凑，你需要在给定物体的外观和构成这个图案必须使用的抽象形状之间反复比较；图片排列，你需要将各种图片存入工作记忆中，并在头脑里将它们重新排列，直到按顺序排好的图片构成一个内容连贯的故事；译码，这是一项完全抽象的任务，主要用来测量信息处理的速度。由于这几项小测验都要求进行某项操作，所以有些人认为这些小测验的得分就构成了操作智商。这些操作都要求测试者当场反应，只用到很少一部分知识储备。

一般智力的另外一个组成部分叫作晶态智力，是指你储存的关于世界本质的信息，以及通过学习得到的有助于你对世界本质做出推论的步骤。韦克斯勒儿童智力测验量表的小测验中最能集中测试晶态智力的是常识、词汇、理解、类同及算术。当然，做算术既要使用储存的或晶态的知识，又要执行某些操作。即使不是全部，这些操作的大部分也是在之前就已经学习过的。韦克斯勒儿童智力测验量表的创建者们，将这几个小测验得到的分数之和称为语言智商，因为绝大多数在测验过程中利用的信息在本质上都与语言相关。操作智商和语言智商得分之和就是总智商，或称为全量表智商。

我们是如何知道一般智力有两种迥然不同的类型的呢？首先，我们称之为操作导向的各项小测验显然更多地利用了推理能力（液态智力）而不是储存的知识（晶态智力），我们称之为与语言有关的各项小测验显然更多地借助知识（包括运用算术解决问题的知识）而非推理能力。不仅如此，各项语言小测验之间的关联度也要高于它们与各项操作小测验之间的关联度，反之亦然。

另外，完成测量液态智力的各项小测验需要的都是执行功能，而执行功能不仅需要额叶皮质的一部分，也就是前额皮质的支持才能得以实现，还需要与前额皮质连接构成网状结构的另外一个区域，也就是前扣带皮层的帮助。前额皮质的损伤会严重破坏完成脑力任务的能力，因为完成脑力任务需要应用工作记忆、注意力控制和抑制控制的执行功能。前额皮质受到严重损伤的人可能根本无法完成瑞文推理测验，根据测验中的表现，他们可能会被视为智力发育迟缓，但他们的晶态智力则有可能完全正常。相反的模式同样存在。患有自闭症的儿童通常都是晶态智力受到了损伤，而液态智力则完全正常，甚至超出常人。

就像我们从有关前额皮质损伤的证据中能够得出的结论那样，脑成像研究也显示，当人们试图解决需要利用大量液态智力的问题时（例如瑞文推理测验或复杂的数学题），前额皮质的活动尤其活跃。

另外一个可以证明智力分为两种类型的证据就是，液态智力和晶态智力的曲线在人的一生中轨迹大不相同。图1–2显示的是这两种智力曲线的理想状态。液态智力在生命的最初几年发展十分迅速，但是在很早的时候就开始衰退。20岁出头的时候液态智力就开始显现出衰退的迹象。数学家，还有一些需要为符号化的、抽象的材料找出新的解决方案的人会发现，他们的能力在30岁左右就开始慢慢衰退了。到了70岁的时候，液态智力下降的幅度已经非常明显了，比1个标准差还多，这就是老年人更难完成拼图游戏和迷宫游戏的原因。而另一方面，晶态智力水平会随着年龄的增加而不断提升，这种提升能够持续到很大的年纪。历史学家以及一些依赖大量信息储备才能有所成就的人会发现，直到50岁的时候他们的能力还在持续增

强。

需要注意的是，上述所有关于液态和晶态两种智力的年龄曲线的内容，在一定程度上还存在争议。争议的详情我就不在这里赘述了，在此只说达成普遍共识的一点，那就是液态智力的衰退早于晶态智力的衰退。

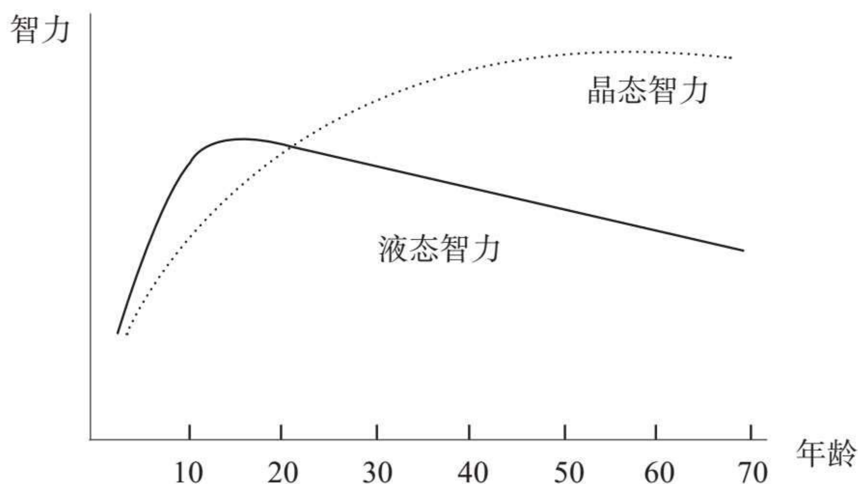


图**1–2** 液态智力和晶态智力的年龄曲线图

资料来源：Cattell, 1987.

前额皮质是大脑所有结构中最早出现衰退的部分，这一点也预示了液态智力早于晶态智力开始衰退。

最后一个可以证明存在两种智力类型的依据就是，执行功能和整体智商是可以分开来遗传的。执行功能在一定程度上遗传自父母。晶态智力，或者说有利于解决问题的知识，也在一定程度上遗传自父母。如果父母拥有较强的执行功能而晶态智力相对较低，那么子女的执行功能也较强，晶态智力也相对不高。

与上了年纪的人相比，液态智力对于年轻人的智力发展更为重要。对于小孩子来说，阅读及数学技能与液态智力之间的相关度要高于其与晶态智力的相关度。与此相反，对于大一点儿的孩子和成年人来说，阅读及数学技能与晶态智力之间的相关度要高于其与液态智力的相关度。后面我会谈到为什么社会经济地位较低的人以及一些少数族裔的人智商相对不高，这一论点在讨论中非常关键。

另一个与液态智力有关的极其重要的事实是，前额皮质与边缘叶之间存在大量互连，而边缘叶与人的情绪和压力有很大关系。当一个人处于情绪激发状态时，前额皮质的活动就会减弱，因此液态智力的功能就会更弱。日复一日，长期饱受压力的折磨可能会永久性地损伤前额皮质的功能。这一信息在我后面谈及穷人及少数族裔的液态智力可变性时，会变得非常关键。

虽然智力是由液态智力和晶态智力共同构成的，但我在讨论时重点将智商值作为一个整体来讨论，只在必要时才会对这两种类型加以区分。

多种多样的智力

智商得分能够预测什么呢？首先，智商得分可以用来预测学习成绩。这一点儿也不奇怪，因为这正是100多年前阿尔弗雷德·比奈发明智力测试的初衷。他想要通过智力测试挑选出那些不太适合正规教育，有可能需要特别的教育方法的孩子。典型智力测试的得分与小学生的学习成绩之间的关联度是0.5。虽然这个数值已经够高了，但这意味着还存在大量智力测试没有测量出的变量，而这些变量也能够预测学习成绩。

智力测试测量的往往是所谓的分析型智力，以及与之相对的实践型智力。分析型的问题通常是人为设计的，对问题的描述很详细，解决问题所需的所有信息都已经包含在对问题的描述中了。分析型的问题只有一个正确答案，通常只有一种特定的策略才能解决，而且和日常生活经验的联系并不紧密，问题本身也没有什么趣味性。这和实践型的问题形成了鲜明的对比。实践型的问题首先需要人们意识到有问题需要解决；对问题的描述往往不够完善，存在多种不同的可能解决方案；经常需要借助日常经验来解决问题；并且通常需要内在动机的参与。

测量实践型智力时，罗伯特·斯滕伯格要求受试者回答在某些情景下应该采取何种应对方式，例如走进一个派对，但参加派对的人你一个也不认识；和几个人合租房子，讨论应该如何分担房租以确保对每个人都公平；为一个你根本不怎么熟悉的人写一封推荐信等诸如此类的问题。

斯滕伯格还提到了第三种类型的智力，即他所称的创造型智力，是指创造、发明或想象事物的能力。他采用命题作文的形式来测试创造型智力水平的高低。例如，他让受试者以“章鱼的腕足”或“第五次机会”为题写一个故事。他还向受试者展示一系列图片，然后要求他们就其中的一张图片编一个故事，或要求受试者为新产品写广告语。

测量分析型智力时，斯滕伯格采用的都是标准化方法，即SAT（美国学术能力评估测试）、ACT（美国大学入学考试）或者智力测试，在测量实践

型和创造型智力时则采用他独创的新奇方法。他发现，测量实践型和创造型智力能够更加准确地预测诸如学习成绩和工作表现等各种结果。有时，这种对于可预测性的提升是十分可观的。实际上，有时这些测试的效果远远好过智力测试。

斯滕伯格利用假想出的三个大学生的例子来进行说明，他的说明非常具有说服力。分析型的爱丽丝非常善于与人讨论各种想法，对他人的产品分析得头头是道。创造型的凯茜对于处理各种观点不太在行，但总有一些独创的有趣的想法，其中一小部分往往最后还真能派上用场。实践型的帕蒂既不善于分析也不是特别有创意，但她总能找到方法并完成任务，她可以用既明智又划算的办法完成某项任务。

你当然希望你的同事们同时具备这三种智力类型，但是在一个团队里一起工作的时候，哪怕只是在其中一种智力类型上很突出的人，都能发挥至关重要的作用。值得我们注意的是，在斯滕伯格对实践型和创造型智力的测试中，少数族裔与多数族裔之间的差距比分析性测试所显示的差距要小得多。这就意味着，这种测试能够使更多的少数族裔进入学校学习，找到工作，避免仅仅由于分析型智力测试的结果不佳就被拒之门外。

霍华德·加德纳认为智力测试测量的只是语言、逻辑数学和空间方面的能力，而忽视了其他方面的“智力”。这些智力包括各种“人际智力”，即类似社会心理学家彼得·萨洛维及其同事研究的“情绪智力”。情绪智力包括能够准确地感知情绪、利用情绪促进思考、理解情绪，以及管理自己及他人的情绪。萨洛维及其同事测量的情绪智力，与智力测试测量的分析型智力几乎毫不相关，但却可以用来预测同龄人及上司评价的诸多维度，例如对人际关系是否敏感，是否善于交际，能否营造积极的工作环境，能否承受压力，以及是否具备领导潜力。有些人也许不愿使用“智力”这个词来描述这些与情感技巧有关的能力，但这是没有道理的。

加德纳谈到的其他类型的智力还包括“音乐智力”和“动觉智力”。一些智力研究专家完全不屑于把“智力”这个词用在这些能力上。但是确实存在与音乐及动觉有关的种种灵感，也确实有音乐及动觉方面的问题需要解决。我个人很乐意将贝多芬的《第七交响曲》和阿尔文·艾利的《启示录》称为天才作品。所以我十分赞成：创造这些作品的本领就是智力的体现。不过，我不会把我个人的偏好强加给那些反对将这些能力称为智力的人。

对于增加的这几种智力类型，加德纳是这样解释的，他指出，在这几种智力类型中大都出现过神童，并且有神经学的证据显示，这几种智力与大脑的不同区域是一一对应的。不管别人称他发现的这几种智力只是技能还是别的什么，很明显，它们与标准的分析型智力不同，测量这几种智力能够预测——或至少在理论上能够预测——人类技能进步的重要方面，而这些

都是标准测试所无法做到的。

动力与成就

最后，某些谁都不会称之为“智力”的特质也会显著影响学业和职业成就。

几十年前，人格心理学家沃尔特·米歇尔对儿童的“延迟满足”能力进行了研究。他将一组来自斯坦福大学幼儿园的儿童分别安排在一个房间内（孩子们以为别人看不到他们，而实际上房间外有一名实验人员在观察他们的一举一动），房间内会放一块曲奇，或一块棉花糖、一件玩具，或者其他有吸引力的物品。研究人员告诉孩子们，他们想什么时候拿到这样东西都可以，只需摇一下铃铛，实验人员就会进来把这样东西拿给他们。不过他们也可以等到实验人员自己进来，如果能等那么久，他们就可以得到两块曲奇，或者两块棉花糖、两件玩具。这一实验测量的就是“延迟满足”的能力。孩子们在摇铃前等待的时间越长，延迟满足能力就越强。

米歇尔等待了十几年，直到这些大多出身中上层家庭的孩子进入高中学习。根据家长们的评价，那些在实验中能够等待更长时间的孩子更能集中精力，做事更有计划性，能更好地面对挫折以及更成熟地面对压力。这些特点在测量学习智力的测试中发挥了积极作用。当年那些幼童的延迟满足能力越强，他们高中时的SAT成绩就越好。拖延的时间长度与SAT语言部分得分之间的关联度是0.42，与数学部分得分之间的关联度是0.57。有可能是越聪明的孩子能够忍耐延迟的时间就越长，不过这似乎并不是唯一的解释。更可信的解释是，越能抵制住诱惑的孩子，长大后越有可能用功读书。后面我们还会注意到，虽然SAT分数与智力测试的分数之间相关度很高，但它并不能完全等同于智商。一些文化群体的SAT成绩比智力测试预测的成绩更优秀——其中的原因很可能与动力有关。

激励因素会影响学业成就的高低，这并不令人惊讶。但动力有时候也许能比智力测试更好地预测学业成绩，却有点儿让人出乎意料。这一点是通过一项极其重要的研究发现的，这项研究在美国东北部地区某大城市的一所磁铁学校[注]内展开。安吉拉·达科沃斯和马丁·塞利格曼将该校的八年级学生作为研究对象，采用多种方法对他们的自律程度进行测试。实验人员不仅询问学生们在多大程度上会产生冲动的言行，还测试学生们对待不同奖励的态度，例如他们是要一个小小的但是马上能够得到的奖励，还是要一个大大的但是要等待一段时间才能得到的奖励。当时他们给孩子们提供了两种选择：立即得到一美元，或等待一星期后得到两美元。他们还向父母和老师询问孩子们约束自己的行为、遵守纪律及控制冲动反应的能力。他们将所有这些测试的得分加在一起，得到测量自律的综合分数，然后比较究竟是这种测试还是标准智力测试，能够更好地预测学习成绩。结果是：

智力测试预测成绩的准确度远远不及动力测试。智力测试分数与成绩之间的关联度仅为0.32，而自律测试分数与成绩之间的关联度则比前者高出一倍多——0.67。自律测试（0.43）比智力测试（0.36）能更好地预测标准的学校测验分数，尽管这两个数值的差距并不大。如果要你为孩子在高智商和强自律性之间做出选择的话，也许选择很强的自律性更明智一些。

虽然达科沃斯和塞利格曼的研究成果很重要，但我们还需要对这一实验进行重复验证。在一所随机挑选的学校，甚至是另外一所磁铁学校里，也许自律测试和智力测试对学业成就的预测效果和在这个实验得到的结果并不相同。尽管如此，这项研究还是能够证明激励因素比智商更能准确地预测一个人未来的成就。

让我们来总结一下我们从上述种种研究中学习到的一些知识。

智商只是智力的一个组成部分。智力测试并不能准确地测量实践型智力和创造型智力，但二者能增加学业成绩和职业成就的可预测性。一旦我们能够进一步改善这两类智力的测量方法，我们也许会发现，实践型智力和创造型智力与智力测试测量的分析型智力是同等重要的。

不论哪种类型的智力，也不论采用哪种方法测量，都只是学业及职业成就的一个预测指标而已。情感技巧和自律，很可能还有其他一些涉及动力和性格的因素，对预测学业及职业成就来说都是很重要的。

除了上述削弱智商重要性的因素以外，我们还要提到一点：只要智力达到一定的水平，大多数雇主似乎并不苛求智商越高越好。相反，他们表示，他们更看重职业道德、可靠度、自律性、毅力、责任感、沟通技巧、团队合作的能力以及应变能力。

所以，智商并不是智力的全部要义。智力，即使赋予了它比智商分数更加广泛的定义，也不是影响学业成绩或职业成就的唯一重要因素。不仅如此，学业上的成功本身也只是预测未来职业成就的一个方面而已。

智商在多大程度上能预测未来？

尽管如此，智商与学习成绩还是和很多结果息息相关，但是很难确定它们之间的因果关系究竟是什么。研究人员通常会确定一些变量，例如个人现在的智商或此前的智商，原生家庭的社会经济地位，童年的生活情况，兄弟姐妹的数量，家里是否有图书馆借阅卡，个人的受教育程度，以及其他一些变量，然后将所有这些变量放入一个多元回归方程式中，用这个方程式预测成年后的社会经济地位、收入或社会病状等。研究人员记录下在排除所有其他变量时（即保持其他变量不变），每个变量对该多元回归方程

式贡献的大小。结果总是证明，在不考虑所有其他变量时，智商对结果而言是非常重要的。但是正如我在关于统计学的附录一中所申明的那样，自变量之间存在着极其复杂的因果关系——天晓得某些自变量会以何种方式影响着其他变量，或者某些自变量会受到甚至尚未测量到的某些未知变量的影响。父母的社会经济地位与子女的学业成绩是有联系的，社会经济地位较高的父母智商也较高，这不仅会影响孩子从父母处遗传的基因，更重要的是这样的父母会更重视子女的抚养及教育，从而促进孩子智力技巧等的发展。所以诸如“智商在职业成就中占百分之几”的说法是最没有统计学依据的。自然所配合的，多元回归不可分开。注

不过，我们可以更精确地对智商决定人生成就的重要性进行定位。政治科学家查尔斯·莫瑞曾经进行过一项研究，研究对象的智商数据都是通过武装部队职业资格测验得到的。武装部队测验是20世纪70年代末进行的美国青年纵向调查的一部分。莫瑞研究了当初的研究对象在成人后的收入和其他社会指标。不过他的研究样本具有高度的选择性——出生在经济状况不算太差（也就是说，家庭收入高于占社会1/4的最低收入家庭）的家庭中的两个子女。他们不能是私生子，至少在他们7岁前父母没有分居或离婚。但两个孩子的智商必须有所区别。一个要在正常范围内——90~109分，一个必须超出这个范围。智商超出正常范围的孩子可以是聪明（110~119分）、非常聪明（120分以上）、迟钝（80~89分）或非常迟钝（80分以下）。

莫瑞能够通过这个被他称为“乌托邦”的样本来确定，智力属于正常范围和超出正常范围的人之间会有多大的差距。他将样本中的个体成年后的收入水平作为最重要的衡量标准。收入当然是与职业成就和社会阶层相关的，所以当我们处理收入这一变量时，可以用收入代表职业成就和社会阶层等其他变量。他同样收集了不少关于样本中的女性是否育有私生子的数据。这一变量也可以代表大量其他的变量，在该研究中是指一些社会功能缺陷，例如坐牢或靠福利救济生活的可能性。

莫瑞发现，即使是在这群家庭稳定、大多数都属于中产阶层的子女中，不同的智商也会产生迥然不同的结果。表1–2显示，如果一个人的兄弟姐妹智商平平，而他（她）却非常聪明，那么他（她）的收入比兄弟姐妹要高出1/3以上——通常（她）的职位相对较高；如果一个人的兄弟姐妹智商正常，而他（她）却非常迟钝，那么他（她）的收入就可能还不及其兄弟姐妹的一半。育有私生子的行为也与智商有着很密切的联系。非常迟钝的女性比其智力正常的姐妹生育私生子的可能性要高出1.5倍。

表1–2 同一稳定的中产阶层家庭内智商不同的两个子女，其智商与收入及女性生育私生子的比例的关系

智商范围	收入（美元）	生育私生子的比例（%）
非常聪明（120分以上）	70 700	2
聪明（110~119分）	60 500	10
对照组（90~109分）	52 700	17
迟钝（80~89分）	39 400	33
非常迟钝（80分以下）	23 600	44

这些分析之所以重要，就在于它们显示出，平均来说，同一个家庭中子女智商不同，其未来的生活状况往往也会截然不同。不仅如此，这些分析无须考虑测试对象原生家庭的社会经济地位带来的影响，因为所有这些比较都是针对同一家庭内部的不同成员的。不过这些分析并不能证明未来的生活状况完全是由智商直接造成的。例如，很有可能受智商影响的教育机会也是因果关系链上的重要一环。实际上，教育很可能会成倍放大智商的作用。另外，智商与同样发挥作用的性格和动力等因素之间，无疑也存在着联系。尽管如此，这些分析结果还是显示了智商的重要性。即使是在稳定的社会阶层较高的同一家庭内部，不同成员的智商也与其未来的生活状况相互关联。

莫瑞研究的智商得分无疑在很大程度上受到遗传的影响。某些孩子在获得父母遗传的过程中比其他兄弟姐妹更走运。莫瑞自己也一直认为智力主要由遗传决定，同时他还认为，正因为这样，智力不容易受到环境因素的影响。但遗传究竟有多大的作用？除了基因，环境扮演了什么样的角色？在下一章中，我将继续讨论有关智力究竟在多大程度上取决于遗传，以及环境究竟能在多大程度上改变人的智力的问题。

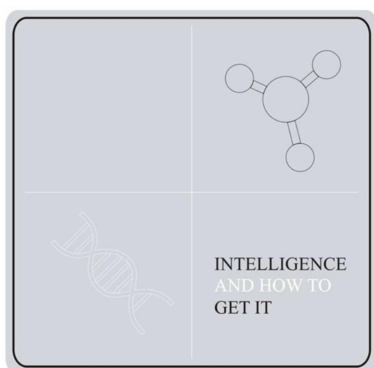
---

1. 如果在某项测试中的成绩很好，那么在与之具有相关性的其他测试中的成绩也往往不错。如果在某项测试中的成绩不佳，很可能在与之具有相关性的其他测试中的成绩也不太理想。

2. 磁铁学校（magnet school），又称磁石学校，是美国一种公立学校，学校提供额外课程以便从广大社区吸引学生。——译者注

3. 作者在这里化用了《圣经·马可福音》中的“所以神所配合的，人不可分开”。

第2章
我们的智力是由什么决定的？遗传基因还是环境？



> ……可以说在智力方面，75%的差异是由遗传变量造成的……25%是由环境差异造成的。
>
> ——阿瑟·詹森（**1969**年）
>
> 孩子们在一个家庭长大，还是在另一个家庭长大，其个性和智力的发展不会有什么差别。
>
> ——桑德拉·斯嘉（**1992**年）

就在不久以前，研究智力的科学家多多少少达成了共识，即智力主要来自遗传。一些科学家至今仍然认为大约75%~85%的智力来自遗传，至少对于成年人来说是这样的。对于生长在同一家庭中的孩子，其共同生活的环境带给他们的影响通常被视为微不足道，尤其是在孩子成年后，这一影响几乎为零。科学家常常相信，或至少从他们的文章中可以看出他们相信，遗传占据了决定性的重要地位，这意味着环境的影响微乎其微，那些以提高智力为宗旨的社会计划注定是不会成功的。

但是如今许多科学家认为智力的遗传度远远低于75%~85%。环境决定论

的支持者估计智力的遗传度是0.50或者更低（不过，正如你随后将会看到的一样，事实上遗传度在不同的人群中有不小的差异）。我同意这些科学家的观点——事实上我怀疑遗传度可能比0.50还要低。

在这一章的第一部分，我将说明为什么早期对智力遗传度的估值如此之高。更重要的是，我将回顾一些收养研究的结果，这些收养研究表明，相对于较低社会阶层的环境，某人在中上阶层的环境中成长可以将智商提高12～18分——这确实是非常大的影响。这为环境影响智力的程度设置了一个很高的上限。最后，我会着重说明，某一特征的遗传性并不能在理论上限制环境对这一特征的影响程度。最后的结论是，环境能够在很大程度上影响智商的高低。如果我们能找到正确的方法改变环境，环境就有可能在更大程度上影响智商。

这一章里有些注释比较长，这是因为我想在不影响大众读者兴趣和理解的同时，回答研究智力遗传性的专家们所关心的种种问题。即使没有注释，这一章也是目前本书中最具专业性的部分。读者们不必在本章花费过多精力。相反，目前请先毫无保留地相信我的断言，即智力绝不是完全取决于基因，环境也可以在很大程度上影响智商的高低。

遗传度、环境与智力

外行们有时认为，某一特征的遗传度是指该特征遗传自父母的程度。这自然而然地就使人们认为，如果智力的遗传度估值为0.80，这就表示一个人智力的80%来自基因。这是完全错误的看法。遗传度绝对不是针对某个个体的，而是指整个群体的遗传度。某一特征的遗传度是指在某个特定人群中，该特征受遗传因素影响而产生变化的程度。与之相对应的是该特征受其他因素的影响产生变化的程度。就智力而言，这些其他因素包括出生前后的生物因素，生物性质的环境因素（例如营养）以及社会因素（例如教育和经历）。我们在这一章里关注的是最有趣的差异来源——遗传和环境，这两者对于生活在同一家庭的孩子们而言是相同的，但对于生活在不同家庭中的孩子来说是有差异的。

不同家庭环境产生的影响，是指包含所有随家庭成员不同而变化的因素，如所处的社会阶层、抚养方式及不同的宗教信仰。不同家庭环境的差异中不包括同一家庭内的环境差异，例如与出生顺序有关的环境差异。因为在一个家庭中只可能有一个孩子是第一个出生的，也只可能有一个孩子是第二个出生的，依此类推。并且我们知道，出生顺序也是影响某些特征的重要因素。同一家庭中的不同孩子在接受学校教育和受同龄人影响方面都存在差异。

当然，家庭之间环境的差异不包括父母的遗传成分在内。因为众所周知，

两个随机选择的家庭，其平均智商差距可能很大，而造成这些差距的绝大部分原因是基因造成的。

那些我称为坚定的遗传决定论者认为，在发达国家中智商的75%~85%来自遗传，而环境以各种方式只起到15%~25%的作用。大多数坚定的遗传决定论者认为，不同家庭之间环境因素的影响（是在史密斯家长大的还是在琼斯家长大的）是微乎其微的，至少在童年以后是这样。同时他们还认为，这种环境因素的影响大多来自家庭内部产生的差异，例如，孩子们所在的学校不同，父母对待他们的方式不同，或是孩子们经历的子宫环境不同。

那么坚定的遗传决定论者是如何得到这些结论的呢？我们来看表2–1。表2–1对大量的研究成果进行了概括，这些研究的内容是不同血缘关系的个体在同一家庭或不同家庭生活时彼此间的智商关联度。遗传决定论者根据分开抚养的同卵双胞胎——具有相同的基因组成——的相关度对遗传度做出了直接估计，这一估值是0.74，注 阿瑟·詹森基本上是基于此计算出遗传对智力的贡献度的。由于被分开抚养的双胞胎所处的环境不同，所以按照逻辑来说，双胞胎之间的相似性就只能来自遗传（以及出生前后的影响。这种影响也许在双胞胎被分开前就已经产生了重要的结果，而遗传决定论者一般认为这种影响是可以忽略不计的）。

表**2–1** 不同血缘关系者被一同抚养或分开抚养时的关联度*

关系	养育方式	关联度
同卵双胞胎	一同抚养	0.85
同卵双胞胎	分开抚养	0.74
异卵双胞胎	一同抚养	0.59
同胞兄弟姐妹	一同抚养	0.46
同胞兄弟姐妹	分开抚养	0.24

（续表）

关系	养育方式	关联度
双亲中值**/孩子	一起生活	0.50
单亲/孩子	一起生活	0.41
单亲/孩子	分开生活	0.24
养父或养母/孩子	一起生活	0.20
被收养的子女	一起生活	0.26

* 这些关联度是对212个不同的研究加以总结并由样本规模加权得到的。
** 双亲中值＝母亲和父亲智商的平均值
资料来源：德夫林、丹尼尔斯，以及罗德（1997年），不包括布沙尔和麦古（2003年）关于一同抚养的被收养儿童的相关度研究。

科学家通常是如何直接对家庭环境差异的影响做出估计的呢？他们对生活在一起但毫无血缘关系的人的智商进行了比较。一种直接做出估计的方法就是确定被收养孩子的智商与其养父母智商之间的关联度。由于他们并没有共同的基因，被收养儿童的智力水平接近其养父母智力的唯一途径就是共处于相同的环境。表2–1的倒数第二行显示这种情况下双方的关联度是0.20。一些科学家认为这是对不同家庭环境对于智力影响程度的正确估值。要得出这一结论的另一种方法是，比较生活在同一家庭中但毫无血缘关系的孩子的智商（表2–1的最后一行）。同样，因为孩子们没有相同的基因，他们智力相似的唯一途径就是共处于相同的环境。此种情况下相关度是0.26，这在不同家庭环境对智商影响的估值中属于偏高的数值。

然而，詹森和其他坚定的遗传决定论者不会接受不同家庭环境的影响系数高达0.20~0.26。他们认为，表2–1中的研究对象大都是孩子，而当研究对象年龄较大一些时，智力与环境之间的关联度就会大幅下降，有时甚至接近于零。这种情况是属实的。例如，对于在一个家庭中被抚养长大的彼此间无血缘关系的孩子们来说，当他们成年时，智力与环境之间的关联度大约变为0.05左右，或者更低。对于环境对成人智力影响较弱这一现象，一般的解释是，随着人们年龄的增长，他们会选择自己的生活环境，而他们对不同环境的偏好主要是受基因影响。早期环境的影响一开始就不大，后来更是一路减弱。这就表示，坚定的遗传决定论者把能够对智力产生影响的环境因素归结为同一家庭内使不同成员产生差异的因素（例如出生顺序），而不是同一家庭内成员所共有的因素，或那些不同家庭之间的差异因素。

让我们来总结一下坚定的遗传决定论者的观点：有3/4或更大比例的智力差异是由遗传造成的；某些智力的差异是由那些家庭内部未共享的环境因素造成的，父母对此也无计可施；成年后几乎没有任何智力差异是由不同家庭之间的环境差异造成的——也就是随机选定的家庭A和家庭B之间的差异。所以你家的特点，或随机选定的琼斯家的特点，对智力几乎没有任何影响。（尽管他们家的收入也许不如你家高，也不像你家一样总读故事给孩子听，孩子们上的学校也比不上你家孩子上的学校，居住的小区设施不健全，宗教信仰也不同。）

此刻，如果你已经为人父母，你可能不禁会想，为什么你还要支付高昂价格以便搬到更贵的、有好学校的社区里去，或者为什么要花费大把银子给孩子买书，矫正牙齿，花时间开车送他们去上小提琴课或陪他们去博物馆，还要使出浑身解数控制自己的脾气，以便给孩子树立一个好榜样？事实上，你根本不必担心对遗传度的估值如此之高，也不用在意对家庭环境的影响估值如此之低。

对智力遗传度的直接估计，是根据分开抚养的同卵双胞胎智力之间的关联度计算出来的，这种做法的前提就是错误地假设双胞胎的成长环境是随机选择的。如果真是这样的话，那么双胞胎之间的成长环境就有大约一个标准差那么大，也就是说，从在谷歌上可以查询到的美国电话号码簿里任意挑选出两个人来，他们之间的差异有多大，双胞胎成长环境之间的差异就有多大。但实际情况并非如此。有可能抚养比利的人（实际上通常是亲戚）和抚养博比的人不会有太大差别。双胞胎各自的成长环境，其相似度会在很大程度上影响两人智力之间的相关性。发展心理学家尤里·布朗芬布伦纳指出，当分开抚养的同卵双胞胎的生长环境高度相似时，两人智力的相关度是0.83～0.91。不过这种相关度不仅反映了他们的遗传基因相同，同时也反映了他们的成长环境是极其相似的。这样的相关度就会导致对遗传度的估值过高。当环境在某种程度上不太相似时(注)，两人智力的相关度是0.26～0.67。由于我们无法确定，在大多数关于分开抚养的双胞胎的研究中，其成长环境究竟有多么不同，我们也就无法确切地从两者间智力的相关度判断出遗传度的高低。

即使不考虑环境的相似度，与从有其他类型血缘关系的人的智力关联度估计出遗传度的方法相比，从同卵双胞胎之间智力相关度得出的遗传度也存在估值过高的嫌疑。这可能是由于虽然分开抚养的同卵双胞胎的成长环境有很大差异，但他们的经历极其类似，因为他们具有十分相像的容貌，或具备其他相同的特征，因此他人往往用相同的方式对待他们。或者也许基因之间存在相互作用，这种相互作用使同卵双胞胎尤为相似，但并不会使除此以外的有血缘关系的人有更加相似的地方。

将遗传度估值为0.75～0.85的第三个错误原因是，双胞胎经历的子宫环境相同。德夫林及其同事坚持认为，同卵双胞胎所处的子宫环境相同，这意味着对遗传度的估值应该减去20个百分点。

后面我会指出对遗传度估值过高的第四个错误原因，即不同的社会阶层其遗传度也不同，而同卵双胞胎研究倾向于从遗传度高的社会阶层中挑选更大量的研究对象。

一旦把这几个方面的错误改正过来，基于同卵双胞胎相关度的遗传度估值就会远远低于詹森及其他坚定的遗传决定论者认为的0.75～0.85。

基因是环境施加影响的触发器

发展心理学家桑德拉·斯嘉和凯瑟琳·麦卡特尼，经济学家威廉·狄更斯，以及哲学家兼智商专家詹姆斯·弗林提出了遗传的作用被高估的另外一个原因，即原本微弱的遗传优势被加以利用后就会转化为巨大的智力优势，原因就在于遗传优势会影响一个人的生命轨迹。让我们举一个关于篮球的例子，比平均身高略高的孩子更有可能打篮球，更有可能享受篮球的乐趣，更有可能经常打篮球，也就更有可能吸引教练的注意，进而进入篮球队打比赛，等等。身高的优势完全依赖于这样的基因条件才能发挥出来。分开抚养的同卵双胞胎很有可能会有非常类似的经历，因为他们身高差不多，所以有可能他们练就的篮球技术也不相上下。但是篮球技术接近并不是因为他们拥有相同的“打篮球的基因”，而是由于他们拥有相同的身高基因，这才导致了他们与篮球有关的经历高度相似。

类似的观点放在智力上也是成立的。例如，如果一个孩子因为遗传而具有相对较强的好奇心，那么家长和老师就有可能鼓励这个孩子去实现各种与智力有关的目标，这个孩子就更有可能发现智力活动的收获很大，就更有可能学习并参与其他的脑力练习。这会使这个孩子变得比不具备好奇心遗传优势的孩子聪明——但这种遗传优势可能很小，并且只有借助环境“倍增器”才能够发挥出来，环境“倍增器”对于实现这一优势是非常关键的。尽管如此，按照遗传度的计算方式，这种基因—环境的相互作用（或基因—环境相互关联，这是遗传学家更偏爱的说法）被归结于遗传性的作用。虽然这并不算错，但这样就低估了环境的作用。

为了更清楚地说明为什么对遗传度的估值抹杀了环境的作用，让我们回到关于篮球的那个例子上。假设一个普通身高的女孩，也许是因为她的哥哥或姐姐打篮球，在车库前的车道上放了一个已经用得很旧的篮筐，这个女孩就开始打篮球。再假设一个高于平均身高的女孩，也许因为家在乡下，邻里四周没有其他孩子可以一起打篮球，这个孩子就失去了接触篮球的机会。来自乡下的这个较高的孩子就不太可能成为篮球队员，而那个身高一

般的孩子则有相当大的可能成为篮球运动员。这两个孩子，一个具备遗传优势却不太擅长打篮球，一个没有遗传优势却成了篮球高手。遗传很重要，如果环境保持不变，遗传也许会在很大程度上决定才能的大小。但是一旦环境介入，就会大大影响——甚至超越——遗传的作用。这一点在估计环境对智力影响的程度大小时尤为重要。不难想象会有很多种方式可以或多或少地使更多人能够学习，而且更愿意学习。

托尔斯泰和领养

现在让我们来看看遗传决定论者是如何评估不同家庭环境的影响的。通过计算被收养儿童智力与其养父母智力间的关联度，以及生活在同一家庭内彼此间毫无血缘关系的孩子智力之间的关联度（通常都是被收养的孩子），遗传决定论者对环境差异的影响做出了估计。正如我们在前面看到的数据，这些情况下的关联度平均来说都很低——0.20~0.25。不过只有当我们假设不同养父母营造的环境之间的差异度与所有家庭环境之间的差异度一样大的时候，这些数字才有意义。而实际情况是，像托尔斯泰所说的“幸福的家庭都是相似的”，收养家庭也都是相似的。

心理学家麦克·斯图米勒证实，在收养家庭中，能够预测智力的因素间的差异程度远远低于所有家庭的整体水平。其中有两个原因。第一，收养家庭的社会经济地位高于非收养家庭。在社会经济地位的阶梯上，最低的一阶上几乎没有收养家庭的身影。第二，收养家庭HOME[注]的得分差异很小。HOME研究者想要通过考察家庭中存在的智力刺激和支持因素的多寡来评估家庭环境。考量的主要指标包括，父母是否经常和孩子谈心，孩子们是否有很多机会接触到书籍和电脑，父母的行为中包含的慈爱多一些还是严厉多一些等等。HOME的评估显示，在这些方面，收养家庭的HOME值远远高出其他家庭的整体水平。实际上，收养家庭的百分位排名是第70位。不仅如此，与整体人口相比，这些变量的变化范围是非常小的。而穷困家庭在HOME测试中得分的差异程度是收养家庭的5倍。也就是说，穷困家庭彼此间的差异程度要比收养家庭彼此间的差异程度大得多。

为什么环境差异小就会造成关联度估值如此之低呢？这是因为如果一个变量变化极小，那么它和与它相关的另一个变量之间的相关度就不可能太高。让我们考虑一下极端情况，变量A根本不变化，那么变量B与变量A的相关度就为零。不论变量B的分值是高还是低，都只能与变量A的同一个分值相联系，所以变量A与变量B之间的关联度只能是零。因此，当不同收养家庭环境之间的差异度被高估时，环境对智力的影响就会被低估。

由于不同收养家庭之间的环境差异被错误地高估为和所有家庭间的环境差异一样大，对不同家庭环境影响的关联度估值就变得非常低。根据斯图米

勒的计算，注如果对环境这一变量的范围进行校正，就会有50%的智力差异要归因于不同家庭环境之间的差异。我们已经知道同一家庭内部环境的差异也会对智商造成很大影响，注这就意味着绝大部分智力差异是由环境造成的。（不过这些结论只在孩子身上成立。我们知道，在一定程度上遗传度会随着年龄的增长而升高，所以我们不得不从斯图米勒关于不同家庭之间环境差异的影响估值上扣除一个未知量。）

智力的遗传度没有固定的数值

那么遗传对智力的贡献究竟有多大呢？实际上，遗传学家表示，根本就不能对遗传度进行单点估值。遗传度取决于特定的人群以及考察遗传度时的特定情况。尤其对于智力来说，人群的性质是非常关键的。心理学家艾瑞克·特克海默及其同事最近证明，遗传度完全取决于人们所处的社会阶层。他们发现，处于社会中上阶层的父母，其智力遗传度为0.70，而较低社会阶层的父母，其智力遗传度仅为0.10。对此比较合理的解释是，社会经济地位较高的父母为孩子智力的发展提供了良好的条件。在这一点上，所有社会经济地位较高的父母之间都没有太大的差别。这种情况下，遗传对智力的贡献就会很大。在极端的情况下，也就是如果所有家庭的环境完全相同，那么遗传就是造成智力差异的唯一因素。

为什么社会经济地位较低的人群，其智力的遗传度如此之低呢？我们通过斯图米勒的研究得知，就能够影响智力的环境因素而言，处于较低社会阶层的家庭之间的环境差异程度，远远大于中等和上层家庭。社会经济地位较低的人，其家庭环境既可能并不逊于最有利于子女发展的中上阶层家庭，也可能方方面面都受到问题的困扰。这就意味着对于这个群体来说，环境因素会对智力起到很大的作用。实际上，环境的影响几乎完全盖过了遗传的作用。注

所以，你花在子女身上的时间、金钱和耐心并没有白费。如果将不同社会阶层的遗传度平均一下，很可能最高不会超过50%。剩下的智力差异大部分是环境因素造成的——同一家庭内部共享的环境，不同家庭之间的环境差异，以及同一家庭内部存在的环境差异。（其余的部分——很小的一部分——是由测量误差造成的。）

同时也请注意，特克海默的研究结果还具有一个极其重要的意义，那就是这样的研究结果也印证了遗传度的估值如此之高，确实是因为被高估了。这是因为对遗传度的估值主要是基于对双胞胎的研究。双胞胎研究大都倾向于选择来自中上阶层的参与者，因为中上阶层的人更易联系，更易被说服参与这样的研究项目。因此，对成年人遗传度的估计是向上偏倚的，而

对不同家庭环境的影响的估计则是向下偏倚的。

斯图米勒关于养父母在促进智力发展方面具备优势的观点，引发了另外一个问题，那就是收养的影响究竟有多大。解决这一问题能为研究家庭环境对智力的贡献提供另外一种方法。如果收养家庭在与能够预测智力的变量有关的研究中堪称示范群体的话，我们难道不该发现被收养孩子的智商要高于根据他们的出身预测出来的智商吗？当然这必须是在环境对智力有很大影响的假设之下。

从收养研究看家庭环境对智力的影响

对遗传度的估计是建立在关联度的基础上的，但正如我们看到的，根据关联度得出的结论也有可能是错误的。我们需要用实验来测试家庭环境的差异会对智力带来多大的影响。从实际情况来看，有大量研究日常收养行为的自然实验可供我们参考。收养家庭环境的好坏是否会对被收养的孩子的智力产生影响，这是我们想要通过这些实验弄清楚的问题。很多自然实验虽然“设计”各不相同，但结论是一致的：在对智力非常有利的环境下成长会对智力发展产生巨大的推动作用。

心理学家克里斯蒂娜·凯普伦和迈克尔·杜伊姆在法国孩子中进行了一项“交叉抚育”的研究。他们跟踪记录了一群出生在不同社会经济地位的家庭里的孩子，这些孩子后来被其他家庭收养。这些收养家庭的社会经济地位有高有低，并且所处社会阶层的差别很明显：以养父的职业作为划分的依据，他们把收养家庭是贫穷的工人阶层家庭（养父是半技工或非技工，受教育时间不超过9年）的孩子，和收养家庭是中上阶层家庭（养父是专业人士或较高级别的管理人员，平均受教育时间为16年）的孩子进行对比。这种安排不仅能够独立评估社会经济地位不同的父母基因对子女智力的影响，还能评估社会经济地位不同的收养家庭对孩子智力的影响。根据研究的结果来看，遗传和与社会阶层相关的环境都会对智力产生巨大的影响。

平均来说，不考虑成长的家庭的社会经济地位，亲生父母社会经济地位较高的孩子比亲生父母社会经济地位较低的孩子的智商高出12分。（我们不知道这一差距有多少是基因造成的，有多少是非基因的因素造成的，如胎儿期、围产期及产后不久的环境因素，但我相信这一差距主要还是由遗传因素造成的。）

这项研究最关键的发现在于，养父母社会经济地位较高的孩子的智商比养父母社会经济地位较低的孩子的智商平均高出12分——不论孩子的亲生父母社会经济地位是高是低，情况都是如此。所以这项研究证明了，在较高社会阶层的环境中成长，能够使孩子的智商远远高于在较低社会阶层的环境中成长的孩子。还有一点也很重要，在较高社会阶层的环境里成长起来

的孩子，其学习成绩也远远好过在较低社会阶层环境里成长起来的孩子。

另一项法国研究采用了不同的“自然实验设计”，研究的是被中上阶层家庭收养的、在社会经济地位较低的家庭中出生的孩子，并将这些孩子与他们未被收养的兄弟姐妹进行比较。被收养的孩子在两项测试中测得的智商分别是107分和111分，而他们未被收养的同胞兄弟姐妹在两项测试中测得的智商平均数为95分。因此我们得出结论，与在较低社会阶层环境中成长相比，在中上阶层环境成长能够为智商贡献12～16分。被收养孩子与其未被收养的兄弟姐妹的学习成绩也有很大的差别。被中上阶层家庭收养的孩子，其考试不及格率只有13%，而他们未被收养的兄弟姐妹的不及格率高达56%。

杜伊姆及其同事在法国进行了另外一项极为重要的自然实验，这项实验的设计与前面的实验都不同，研究的是受到虐待的低智商儿童。这些孩子在四五岁被收养时进行一次智力测试，然后等到他们14岁时再接受一次智力测试。研究人员有意寻找那些被不同社会经济地位的家庭收养的孩子。年龄较小时，孩子们的智商在61～85分——这是智力测试人员认定的智力明显迟缓和智力迟钝—正常的范围。收养这些孩子的家庭被划分为贫穷（非技术工人）、中低和中等阶层（低层或中层管理者、小业主以及技术工人），或中上阶层（专业人士以及高层管理者）。收养对于智力的作用很大，平均使孩子们的智商增长了14分，但是收养家庭的社会阶层使增长的幅度产生了很大的差别。被较低阶层家庭收养的孩子智商增长了8分，被中等阶层家庭收养的孩子智商增长了16分，被中上阶层家庭收养的孩子智商增长了差不多20分。由此估计，与在较低阶层家庭长大相比，在中上阶层家庭长大能够将智商提高12分。

为了能够确切地推断养父母的社会经济地位对孩子的智商产生的影响，研究人员没有对哪些家庭收养哪些孩子进行刻意的安排。也就是说，智商相对较低的孩子与智商相对较高的孩子都有可能被社会经济地位较低的家庭收养。这项研究得出的结论与前面提到的两项法国研究的结论相同：一个孩子在较低阶层环境里长大或是在中上阶层环境里长大，智商会有12分的差别。请注意，这项研究表明，与在社会经济地位较低的家庭成长相比，在相对适中的中低或中等阶层家庭成长能够使智商得到很大的提高，也就是8分。另外需要注意的是，这项研究低估了在社会经济地位较高的家庭（相较于社会阶层较低的家庭）成长所带来的影响。正如斯图米勒证实的那样，即使是社会经济地位较低的收养家庭，其某些养育方法也能促进智力的发展。

一项对截至2005年的所有设计合理的收养研究的回顾发现，和仍生活在出生家庭内（一般社会经济地位较低）的兄弟姐妹相比，被中等或中上阶层家庭收养对被收养孩子的影响是1.17个标准差，也就是说在中上阶层家庭

中被抚养长大的孩子比在较低阶层家庭中被抚养长大的孩子具备18分的智商优势。这次回顾还对遗传因素以及胎儿期、围产期和产后不久的因素产生的影响做出了评估。这一评估是通过在中等和中上阶层家庭的亲生子女和该家庭收养的子女之间进行比较得到的。碰巧两者的差距也是12分，与凯普伦和杜伊姆发现的数值相同。

从这些发现中我们得出了一个重要结论，即出生在较低阶层家庭中的孩子，即使根据其出身预测的智商不高，只要在足够丰富的认知环境里成长，他们的智商就能够大幅提高。

不过对收养研究的检测显示，学习成绩方面的情况却不太乐观。被收养的孩子在学习成绩的测量中，仅仅比他们未被收养的同胞兄弟姐妹高出0.55个标准差。此外，他们在学业成绩的测量中得分仅比全部样本的整体水平低0.25个标准差，和同班同学比起来差得还要更少一些。

在结束有关收养的话题前，我想要指出一点，通常亲生父母与其后代智力之间的关联度，要远远高于养父母与其收养子女智力之间的关联度，坚定的遗传决定论者通常凭借这一点就坚持认为智力首要取决于遗传，而与环境的关系不大。遗传决定论者相信，被收养孩子成长的环境对于其智力没有任何影响，因为收养环境之间的差异不会造成智商上的差异。现在我们可以看出这个结论是多么的荒谬。收养家庭主要是由中等或中上阶层家庭构成，因此家庭环境大都类似。即使是那些社会经济地位较低的收养家庭，其某些养育因素方面的得分也很高，而这些养育因素都预示着孩子的智商可能较高。由于收养家庭之间的差异相对很小，因此养父母与其收养的孩子智力之间的关联度就不可能太高。在那些能决定智力水平高低的因素方面，养父母营造的环境基本没有太大差异——并且如果差异很小，关联度就不可能太大。但是总体上，收养环境与低社会经济地位的环境之间的差异很大，这就造成了智力的巨大差异。所以，养父母与其收养的孩子智力之间的关联度较低，只不过是遗传决定论者使用的障眼法而已——这与收养家庭会对其收养孩子的智力产生很大影响的事实，实际上一点儿也不冲突。

最后，既然赫恩斯坦和莫瑞在《钟形曲线》一书中宣称收养对智商影响的平均值是6分——这是已经达成的“共识”——那么就很有必要指出，他们的证据来自查尔斯·洛库尔托的一份回顾报告，而实际上洛库尔托给出的收养影响平均值是12分。

家庭环境的差异对智力没有任何影响，这恐怕可以算得上是被高智商人士所接受的最不寻常的看法之一了。朱迪思·瑞奇·哈里斯写出了非常有趣的畅销书《教养的迷思》，全书的前提就是假设不同家庭之间的差异对于智

力的影响为零。史蒂芬·平克在其精彩的著作《白纸一张》中也坚持这样的观点。在畅销书《魔鬼经济学》中，史蒂文·列维特和斯蒂芬·都伯纳明确表示，收养对智力的影响微乎其微："已有研究证实，亲生父母的智力对一个孩子学习能力的影响，比其养父母的智力对孩子的影响要大得多。"（我真希望我能避免与这些相信奇怪看法的人为伍，但不幸的是，曾经有很多年，我都持有遗传决定论者关于家庭环境根本不重要的主张，虽然我非常困惑不解。）

通过前面看到的有关遗传基因与环境影响智力的证据，我们了解到一些有关社会阶层和智力间关系的重要结论。与来自典型的社会经济地位较低家庭的孩子相比，来自专业人士家庭和中等阶层家庭的孩子的经历使其智商要高得多，考试不及格率也低得多。另外，按社会经济地位较低的家庭共有的环境因素限制智力发展的程度，我们可以推算出一个数值，或者至少是一个范围：12~18分。不论遗传度究竟是高还是低，都不能推翻这个事实。所以我们知道，在理论上，种种干预措施不但能够非常有效地提高穷孩子的智商，还可以大大降低社会阶层较低的孩子考试不及格的比率，最低估计为0.5个标准差，最高估计要高得多——1个标准差，或者降至由中等阶层家庭亲生父母抚养长大的孩子的考试不及格率水平。

还需要注意的是，不是只有社会经济地位较低家庭中的孩子的智商才会受到影响。一项研究调查了一群被收养的白人孩子的智商，这些孩子的生母智商平平，收养他们的主要是中等或中上阶层的家庭。被收养相对较晚的孩子童年期的平均智商是112分，被收养相对较早的孩子的平均智商是117分。这项研究表明，即使是那些在一般环境中长大将会智商平平的孩子，如果让他们在非常有利的环境中成长，其智商也会得到极大的提高。同样，凯普伦和杜伊姆进行的交叉抚养的研究表明，如果在贫困中长大，即使是原本来自中上阶层家庭的孩子，其智商也会下降，这样的智商损失大概是12分。所以，不只是出生在社会经济地位较低家庭的孩子，有可能因环境导致智力显著变化，所有的孩子都有这个可能。

遗传度与智力可变度无关

现在，我可以向智力高度遗传意味着环境对智力的影响微乎其微这一观念发起最后一击了。智力可变度可能的变化范围不会受到遗传度的限制。这一点太重要了，我需要再重复一遍，智力可变度可能的变化范围不会受到遗传度的限制。所有遗传学家都接受这一原则，遗传决定论者经常承认这一原则，但同时又持续主张一些似乎事实上遗传度确实限制了可变度变化范围的言论。

为了理解为什么可变度不会受到遗传度的限制，让我们来思考两个事实：

（1）身高的遗传度大约是0.85~0.90；（2）世界上有一些国家，在一代人或更短的时间里平均身高增加了1个标准差或更多。1965—2005年，韩国13岁男孩的平均身高增加了不止7英寸，[注]也就是2.40个标准差。1965年身高正常的男孩，在2005年看来简直就是小矮人。遗传很难在短短40年的时间里使身高实现如此大的变化。我们看到，在两三代人经历过的岁月里，世界上许多地方的人的身高都有所增加，这显然是某种环境变化造成的结果，很可能是营养环境的变化。

或者我们可以考虑一个更加极端的例子：假设遗传度为100%，并且环境的影响也非常大。我们可以随机将谷物的种子抛撒到肥沃或贫瘠的土壤里。想象一下在肥沃的土壤里谷物的遗传度为100%，在贫瘠的土壤里遗传度也是100%。尽管如此，两组谷物的平均高度仍会大不相同，这种差异完全是由环境因素造成的。

这些例子应该很清楚地说明了，特定人群的某一特征的遗传度在理论上不会限制环境因素改变该特征的程度。这很好，因为正如你将在下一章看到的那样，除了受到遗传度的适度影响，智力同样在很大程度上取决于环境因素——也就是在过去80年里发生在学校里及社会上的种种变化。

---

1. 如果你对关联度有一定了解，但不太了解有关遗传度的研究，你可能就会问，为了求得遗传决定智力的百分比，为什么没有对关联度0.74进行平方。答案是，遗传关联度本身就是平方后得到的数值。

2. 泰勒在1980年也发现，环境的相似会对同卵双胞胎智力之间的关联度造成极大影响。但是布沙尔在1983年使用了不同于泰勒使用的智力测试进行了测量，得出的结论却是：环境的相似并不会对同卵双胞胎智力之间的关联度造成多大的影响。

3. HOME，即Home Observation for Measurement of the Environment，环境监测家庭观察法，一种评估家庭环境的方法。

4. 行为遗传学家马特·麦古及其同事（2007年）在研究了收养家庭和非收养家庭后发现，和非收养家庭相比，收养家庭在社会经济地位及面对的社会问题这两个方面的差异并不大。另外，他们还发现，如果对环境这一变量的范围进行修正，被同一家庭收养的孩子智力之间的关联度，并不比没有修正时得到的关联度高。尽管如此，我们还是要在以下两个事实的基础上来看待这两个发现：第一，他们研究中的所有家庭，包括非收养家庭，都有两位青少年住在家里，这样的家庭更加稳定，社会经济地位也要稍高于一般的家庭；第二，非收养家庭中拒绝参与研究的母亲，其受教育程度

要远远低于非收养家庭中同意参与研究的母亲。因此，参与研究的非收养家庭社会经济地位更高，并且很可能比一般的非收养家庭更加稳定。正如我们在后面将看到的那样，这些社会经济地位较高的家庭的遗传度，要远远超出全部人口的整体水平。

5. 斯图米勒（1999年）证实，正如错误地估计家庭内环境差异的大小会导致遗传对智力的影响程度被高估一样，错误地估计家庭内环境差异的大小也同样会导致家庭内环境差异对智力的影响程度被高估。

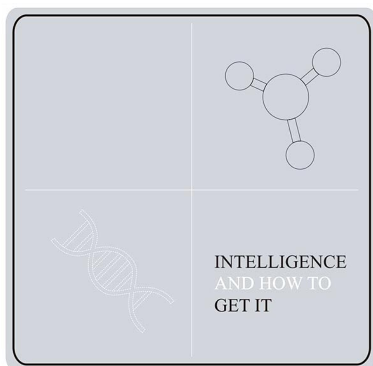
6. 特克海默（我们曾经当面交流过）认为，很多这样的家庭不能够提供足够有利的环境，使得基因的作用无法发挥出来，人与人之间基因的差异也无法显现出来。和社会经济地位较低的家庭间环境存在巨大差异相比，这更可能是造成这个群体遗传度很低的原因。

7. 遗传决定论者或许会抱怨伊詹多恩对收养家庭作用的估计太高，因为这种估计是以幼童为基础的，相对于年龄较小的人来说，遗传度对于年龄较大的人的作用更大——据推测，这是因为随着年龄的增长，人们能够更自由地选择他们想要的环境，而那些具有高智商基因的人们就会选择能使他们变得更聪明的环境。但是伊詹多恩及其同事（2005年）发现，测试时的年龄——12岁以下或13～18岁——不会影响对收养作用的评估。有关年龄和遗传度关系的研究得出的结论也是如此。从童年期一直到青少年末期，人的遗传度都是恒定的。

8. 正如本章中指出的那样，有些人认为，被同一家庭收养的不同孩子在成年后智力之间的相关度极低，这就能够证明，同一家庭内部或许存在环境差异，但这种差异在孩子们成年后就完全不存在了，因为那时他们有能力选择自己想要的环境，他们的遗传潜力也就可以全部施展出来了。他们还认为，拥有高智商基因的人会选择更有利的环境，而拥有普通智商基因的人会选择不那么有利的环境，所以童年时的环境和现在的智力水平没有什么关系。不过，唯一可以支持这一结论的、有关关联度的研究，就是一项纵向研究，这项研究关注的是从童年期到成年期这段时间内被同一家庭收养的兄弟姐妹之间的关联度——样本一直是研究开始时的那群人。这样一来就解决了一个问题，那就是和研究儿童样本相比，研究成年样本时，研究人员研究的往往是社会经济地位较高的家庭里收养的孩子。实际上，似乎只有两项纵向研究，虽然这两项研究都显示了从童年期到成年期关联度是越来越小的，不过这些样本规模都很小，并且童年期时的相关度与成年期时的相关度之间的差距并不大。

9. 1英寸≈2.54厘米。——编者注

第3章 学校教育能让我们变得更聪明吗？



……即使是再完美的教育制度，也不会使智力处于低水平的孩子的表现取得任何进步……

——查尔斯·莫瑞（2007年）

……一个人在瑞文推理测验中的得分总和，可以用来衡量这个人智力水平的高低，这与国籍或教育经历没有任何关系。

——瑞文和考特（1975年）

极端的遗传决定论者认为，环境对智力不会有任何影响，智商高低都是遗传基因早就决定了的。由这一观点可以得到两个重要推论：学校对智商的高低没有什么影响；除非有基因工程，否则整个人类的智力不可能发生大的改变。

这些预言事关重大，结果却被证明完完全全是错误的，这种情况恐怕在行为科学中也不多见吧。

学校教育会让我们变得更聪明吗？

一些心理学家已经非常明确地表示，他们坚持认为教育与智力没有太大联系。教育只是告诉人们具体的事实和固定的方法，而不能提高面对不熟悉的情况时未经训练的解决问题的能力。赫恩斯坦和莫瑞在《钟形曲线》一书中的看法代表了不少智力研究理论家的观点。他们承认，受教育时间的长短与智商的高低确实存在一定的联系，但这种联系主要是因为更聪明的人愿意在学校里待更长的时间。更聪明的人更喜欢学校，所以越来越喜欢留在学校里，因此他们受的教育就越来越多。这一结论是建立在多元回归分析上的，但正如我在第1章中指出的那样，多元回归分析通常不能解决因果关系的问题。如果一个人无法接受教育，那他会怎么样？他会不会因此不如接受教育后那样聪明？很多实验都是围绕着这些问题进行的，事实证明实验的结果都是一样的。

进行这方面研究的社会科学家都是心地善良的人。因为这些实验都是在自然状态下进行的，实验对象都是出于各种各样的客观原因有一段时间未能接受学校教育，并不是因为要实验而刻意安排的。这就排除了赫恩斯坦和莫瑞所说的情形，即受教育时间的长短与智商的高低有联系，只是因为更聪明的人愿意在学校里待更长的时间。发展心理学家斯蒂芬·塞西和温迪·威廉姆斯曾经详细地描述了这些研究。

其中一项自然实验就是暑假。孩子们在暑假期间不上学，这使得他们的智力和学习技能有所下降。下降程度虽然不大，却是真实存在的。对于高年级的学生和社会经济地位较低的孩子来说，这种退步很明显，尤其在数学方面。事实上，社会经济地位较低的孩子与社会经济地位较高的孩子在学习成绩上存在差距的主要原因就是，社会经济地位较低的孩子在暑假里学习退步得更多。

有关学校教育影响智力的研究早在1923年就开始了。心理学家休·戈登研究了伦敦一些流动人口的子女的智商。这些流动人口包括船夫和吉卜赛人，他们的孩子很少上学，或根本不上学。在本该开始接受学校教育的年龄，这些孩子的智商都在低—正常范围的低水平区域内，但此后却持续下降。4~6岁孩子的平均智商是90分——接近正常范围的下限，而年龄最大的孩子（12~22岁）的平均智商是60分，远远低于人们规定的智力缺陷的下限。而上学的孩子平均智商则没有下降。所以研究结果表明，学校教育对于维持孩子的智力水平是非常必要的。

20世纪初，美国边远地区的一些孩子接受到的教育很少，有些根本没有接受过教育。这促成了另外一项早期的自然实验。在围绕着蓝岭山脉的那些“凹地”里，有一些最初的苏格兰—爱尔兰和英格兰移民的子女。这些移民的祖辈在19世纪时将土地立契转让给了德国移民，然后动身前往遥远的高地。这些孩子大都没有机会上学，也接触不到报纸或电影。他们接受了几项测量操作智力的测试（例如积木图案），这些测试都不要求具备读写

技能。结果显示，这些孩子年龄越大，智商越低。但其中一个居住地的孩子接受了一些学校教育，他们的智商并没有随着年龄的增长而下降。

第二次世界大战为另外一项自然实验提供了机会。由于纳粹分子的包围，荷兰孩子上学的时间被推迟了几年。这些孩子的平均智商比包围解除后适龄入学儿童的智商低7分。

20世纪中期，在南非的不少村子里，由于没有老师，一些印第安人后裔的子女上学时间被迫推迟了一段时间，最多的长达4年之久。和邻近村子里有机会上学的孩子相比，每晚一年上学，这些孩子的智商就平均下降5分。即使在学校里学习了几年之后，这些孩子的智商也始终没能赶上其他孩子。

另外一项关于推迟接受学校教育的研究考查了弗吉尼亚爱德华王子郡的一些黑人儿童。1959—1964年，为了避免种族融合，当地关闭了所有公立学校。在此期间没有上学的孩子，每少上一年学，智商平均下降6分。

过早辍学同样不利于智力发展。两组不同的瑞典心理学家获取了几万个男孩的智商数据，这些孩子都是被随机挑选出来的，接受智力测试时是13岁。研究人员挑选出在13岁时智商、社会经济地位和学习成绩都相同的孩子。所有这些孩子在18岁为了申请入伍再次接受了智力测试。心理学家发现，那些在13岁时智商、社会经济地位和学习成绩都相同的孩子，每少上一年学，智商就平均下降2分。那些14岁就退学了的孩子，智商平均下降了8分——等于0.5个标准差。请注意，这些研究都证明了，并不是因为聪明的孩子在学校里待得更久就变得更聪明，而是因为（不考虑13岁时智商是多少）留在学校里学习使智商有了大幅提高。（或者换句话说，离开学校使智商大幅降低。由于平均智商被人为地设定为100分，我们不能确定究竟是辍学的孩子智商有所下降，还是留在学校里继续学习的孩子智商有所增加，或两者兼有。）

孩子们若想在某年开始上学，其出生日期必须在某个特定日期之前才可以。这为证明上学能使孩子变聪明提供了绝佳的机会。例如，许多地区都将准许入学的出生截止日期定在9月的某一天。为了便于说明，我们以9月15日为例。一个出生在9月16日的孩子不得不比出生在9月15日的孩子多等待一年才能进入学前班。这就非常巧妙地构成了一项自然实验，为我们研究学校对智力的影响提供了机会。有的孩子拥有比其他孩子几乎大一岁的优势，有的孩子拥有早上一年学的优势。我们比较这两类孩子的智商后就会发现究竟哪一点更重要——年长一岁还是早上一年学。西里尔·伯特爵士和瑞文，以及他们现在的追随者，都非常明确地预言：对于年幼的孩子来说，年龄差一岁意味着智商差很多，而早上一年学根本不会对智力有任何影响。（确切地讲，瑞文会说，对于类似瑞文推理测验这样的智力测试

来说，一年的学校教育根本不会使智力有任何提高，因为瑞文推理测验测量的是未受文化影响的纯粹的液态智力。）

而实际上，德国和以色列的研究都发现，一年的学校教育等于年龄增长两岁。

对于未接受过西式教育或根本未接受过教育的孩子来说，西式教育也会对其智力产生很大的影响。西式教育有助于增强学生的记忆力，包括智力专家经常宣称的那种根本不会受学习影响的记忆力，例如通过数字广度（记忆口述的数字的能力）和译码测试（按照所给的例子推算出符号与形状或数字的对应关系）测量的记忆力。一些非洲青少年刚刚接受了三个月的西式教育之后，其完成智力测试中出现的一系列空间感知任务的能力就提高了多达0.7个标准差。这些任务包括一些测量操作智力和液态智力的测试，例如积木图案、图形记忆、图画描述——而智力研究人员通常都认为这些任务测量的是未经训练、未受教育影响的智力。

学校的教学内容中就包括那些出现在综合智力测试中的内容，例如《哈姆雷特》的作者是谁、水的组成部分是什么，以及词汇和数学运算。有些智力理论家却仍然怀疑学校能否提高人的智力水平，这实在让人难以理解。传统的智力理论家更难以接受的是，学校同样可以提高人们解决问题的能力。而这些问题，例如出现在瑞文推理测验中的问题，都被视为与文化无关。人人都接触过圆形、方形和三角形，所以很多智力理论家就认为这些完全抽象的测试与学校教育毫无关系。正如你将要看到的那样，这种看法会被证明是无稽之谈。

我们比我们的曾祖父母聪明吗？

既然学校教育可以使我们变得聪明起来，而且现在我们比100年前的人受到的教育要多得多，我们是不是自然应该比我们的曾祖父母要聪明得多呢？1900年，美国人接受学校教育的平均时间是7年，而有1/4的人口接受学校教育的时间不足4年。如今接受学校教育的平均时间是14年，也就是高中毕业后再继续学习两年，并且绝大多数人都读完了高中。

如果你知道平均智商在几乎一个世纪里都是100分，你也许会认为教育对智力没有任何影响。但其实我们在设计智力测试时就将平均智商定义为100分，所以平均数100这个常数实际上并不能反映智力随着时间的推移所发生的变化。要确定现在的人在智力测试中的得分是否更高，就要让他们参加前人进行过的测试，然后将前人的表现与现在的人的表现进行比较。但这需要我们不断重新编写智力测试。如果年年的测试题目都相同，智商得分一定会越来越高。为了将平均智商保持在100分不变，就需要编写新的测试题目，不断增加测试的难度。

所以，就智力测试所测量的种种技能来说，实际上每一年人们的能力都是有所提高的。主要的智力测试——例如，韦克斯勒儿童智力测验量表，韦克斯勒标准成人智力测验量表，斯坦福–比奈智力测验量表——显示，1947—2002年的55年间，智商每年增加了差不多1/3分，如果以30年为一代人来计算的话，这相当于美国每一代人的智商都总共增加了9分。这一智商的增加被詹姆斯·弗林记录了下来，并被命名（不是他自己命名的）为弗林效应。对这种变化进行调查后发现，所有发达国家中都出现了智商快速增长的现象。有些国家比美国增长得要慢一些，有些国家则比美国增长得还要快。

这种大幅增长是什么原因造成的呢？下面我会严格按照弗林在一本重要著作中给出的最新解释，来说明智商快速增长的原因。

有人怀疑，智商之所以增长或许是因为人们的应试能力增强了——更加熟悉标准化的书面测试。这种解释不太能服众。智商的提高至少从1917年就开始了。从1917年到第二次世界大战征兵开始，18岁的年轻人参加军队智力测试得到的分数提高了12～14分，那个时期人们恐怕不太可能接触到标准化测试。从几乎没有什么人接触过标准化测试，到最近几十年人人都熟悉标准化测试模式，这段时期内智商差不多都在稳步提高。不管怎样，如果从来没有参加过测试的人不断重复参加智力测试，他们的测试得分一定会有所增加，但不会增加很多。而我们看到的事实是，在1947—2002年这段时期，智商分数提高了18分之多。

智商的增长和营养有关系吗？恐怕也没有太大的联系。虽然如今在世界上的某些地区，或许甚至在第二次世界大战前的美国和欧洲，营养不良毫无疑问对智商产生了不利的影响，但时至今日，没有证据显示严重的营养不良会影响很多人的智力发展。如今发达国家中大部分营养不良的情况都发生在胎儿期或产后不久，这种营养不良的情况在最近几十年里有所缓解。不过，一直以来都有人认为，营养不良并不会影响人口的整体智力水平。每有一个孩子因为围产期营养更充足而提高了智力水平，就会有另一个孩子从死亡的边缘被拯救过来而智力受损。不管怎样，几乎可以肯定的是，最近几十年里发达国家中很少有孩子受到营养不良的影响。

还有一点也可以说明智商的增长与营养的改善没有关系。智商得分在智商的数值范围内是匀速递增的。智商数值位于前1/3的人的智商和智商数值位于后1/3的人的智商，提高幅度是一样的，在距离现在较近的任何时代，这一部分人都不会在营养方面处于劣势。这种递增也恰好证明了查尔斯·莫瑞是错误的，因为他曾经说过，没有什么方法能使智商值处于后半段人的智商得到提高。

那么这种智商的提高究竟说明了什么呢？我们还得先来弄清楚：在55年的

时间里智商提高18分，以及在之前的30年里智商提高了9分甚至更多，究竟具有怎样的意义。我们假设1947年时智商的实际数值是100分。智商为100分的人通常都是技术工人，不承担太多责任的办公室职员以及售货员。这些人即使有钱上大学，恐怕也很难顺利毕业。如果智商处于平均水平的人的孙辈参加相同的智力测试，得到的平均智商会达到118分。智商为118分的人不但可以出色地完成大学的学习，如果愿意，还有可能继续攻读研究生，并成为诸如医生或律师这样的专业人士或者高级经理、成功的企业家。以这样的标准来看，人有没有可能变得更聪明呢？

或者让我们来反着推算。我们假设2002年人们平均智商的实际数值是100分。如果使用的智力测试与今天智商处于平均水平的人参加的智力测试相同，这个人的祖父母的平均智商就是82分。那么这些祖父母就不太可能胜任需要承担一定责任的办公室工作，也不太可能符合对大多数技术工人的工作要求，甚至能否高中毕业都很难说。

或者让我们再倒退到1917年。如果使用今天的智力测试，智商处于平均水平的人的曾祖父，其测得的智商只有73分！这样的智商恐怕连个技术工人的工作都很难找到，也根本不可能从高中毕业。按今天的标准来衡量，那时一半的人口都属于智障！

显然，这样的推算出现了问题。我们没有那么聪明，而上几代的人也没有那么笨。

我们究竟在哪些方面变得更聪明了？

不过，我们知道我们一定是变得比以前聪明了，因为学校教育能使我们变得更聪明，如今我们接受教育的时间也比祖辈们要多得多。那么我们究竟聪明了多少？在哪些方面更聪明了呢？

为了更好地回答这个问题，让我们来看一下韦克斯勒智力测验量表以及使用最广泛的“与文化无关”的测试——也就是瑞文推理测验的得分情况。图3–1显示的是1947—2002年间几项主要的智力测试得分的变化情况，包括瑞文推理测验、韦克斯勒智力测验量表（该测试已被用于测量数以百万计的6～16岁的孩子的智商）、韦克斯勒智力测验量表中测量液态智力的5项“操作”小测验（图画补缺、积木图案、图形拼凑、图片排列以及译码测试）、韦克斯勒智力测验量表中测量晶态智力的两项小测验（类同和理解）、韦克斯勒智力测验量表中测量晶态智力的其他三项小测验（信息、词汇和算术）。请注意，为了便于比较，所有这些测试和小测验所参考的平均值都是100分。

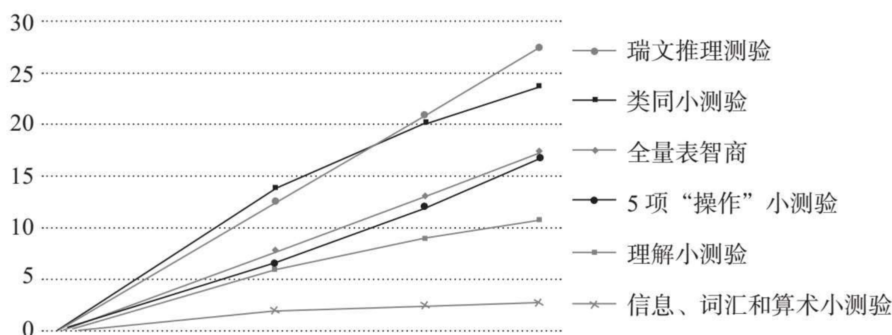


图**3–1** 美国**1947—2002**年韦克斯勒全面智力测试、
韦克斯勒小测验以及瑞文推理测验的得分情况

资料来源：Flynn, 2007, p.8.

在这张图中，代表不同测试的几条线交织在一起。瑞文推理测验以及部分韦克斯勒儿童智力测验量表小测验的分数大幅增加，而剩下的一些测试分数几乎没有太大变化。

让我们首先来讨论据说“与文化无关”的瑞文推理测验（见图1–1）。这项测试测量的是未经训练过的智力，也就是不易受文化或学校教育影响的智力。该测试显示，平均智商的增长超过了28分！也就是说，1947年智商为100分的人，现在他的孙辈的智商几乎达到该测试定义的天才的水平。我们绝对可以肯定一点，人的实际智商，也就是广泛的解决问题的能力等，不可能有如此大的增长。我们还可以肯定一点，瑞文推理测验绝不是和文化无关的测试，文化已经完完全全渗透到了瑞文推理测验当中。这是因为，基因不可能在这段时期内发生如此巨大的变化，任何可能影响智力的营养标准或其他生物因素，也不可能使智商实现这样大幅度的增长。

那么为什么人们在瑞文推理测验中的得分提高了这么多呢？我们不知道答案，不过我们可以理性地猜测一下。发展心理学家克兰西·布莱尔及其同事证实，在幼儿园和小学就开始教授的数学，其教学内容早就发生了变化，原来只教孩子们计数和算术运算，现在则多用几何图形和非常直观的形式来代表物体，然后要求孩子们指出其中的变化规律。图3–2就是近几十年来孩子们经常接触到的一个直观显示的例子。你可以看到这对于解决实际中的瑞文式的问题是多么有帮助。发展心理学家温迪·威廉姆斯已经证实，与以往相比，现在的课堂教学在总体上更加强调解决抽象的感性问题的能力。布莱尔及其同事还证实了一点，与以往相比，现在的数学问题更要求学生具备解决一系列错综复杂问题的能力，这种能力对于解决瑞文式问题所需的工作记忆能力是至关重要的。

运用批判性思维

运用批判性思维

按照规律完成每一行最后一个图案。

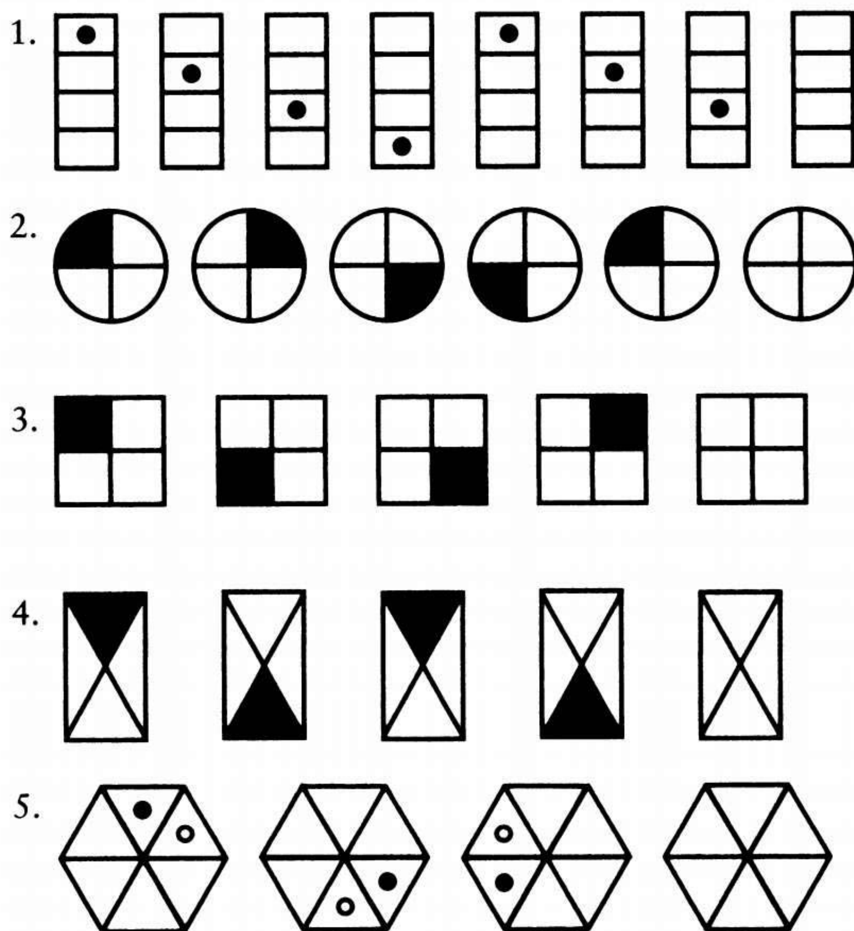


图3–2 现代幼儿使用的算术课本中的一个直观问题的例子

资料来源：Eicholz, R. 1991, p.56.

根据图3–2显示，不仅仅是瑞文式问题，几乎所有与操作智力和液态智力类型有关的技能都得到了提高。这得益于文化不断地向更加直观的刺激形式靠拢——课本、电视、孩子们的游戏书和电脑（包括电脑游戏）。例如，韦克斯勒儿童智力测验量表所有测试操作智力的小测验都含有大量画面成分，并且大部分测试题都需要在工作记忆中完成复合操作才能够解

决。

我们完全有理由相信，这些直观练习实际上会提高与液态智力有关的技能以及构成液态智力基础的各种执行功能，包括工作记忆和注意力的控制。例如，研究人员已经证实，和不玩电子游戏的人相比，玩电子游戏的人不但可以同时顾及更多的事情，还能够更有效地忽略不相关的刺激，看问题的视野也更加开阔。为了确定研究人员从电子游戏玩家身上观察到的不只是自我选择效应（那些液态智力较高的人本身就更有可能玩电子游戏），研究人员让不玩电子游戏的人学习玩一个游戏——“荣誉勋章”——研究人员认为玩这个游戏可以锻炼玩家控制注意力的能力。同时，研究人员让其他不玩电子游戏的人学习一个他们认为不会锻炼玩家控制注意力能力的游戏，也就是“俄罗斯方块”。研究对象连续10天每天玩一个小时游戏。10天后，“荣誉勋章”的玩家比“俄罗斯方块”的玩家在完成注意力控制方面的任务时表现得更好。

神经科学家已经证实，可以利用电脑游戏来训练幼儿的执行功能，而执行功能是液态智力的基础。研究人员萝莎里奥·罗达及其同事将研究的重点放在注意力控制上。他们让4岁的孩子在5天之内在电脑上做一系列的练习。例如，孩子们要使用操纵杆来控制一只卡通猫，使它待在草地上而不能进入泥地里。孩子们还做了一些预测练习（例如学习如何预测一只游过池塘的鸭子的动作），他们同时还完成了一个任务，这个任务要求他们记住不同卡通形象的特点。另外，孩子们还要从两个数列中挑出总和较大的那一列。研究人员设计的冲突在于，总和较大的那一列的数字每一个的数值都很小。例如，一列数字是由9个4组成的，另一列数字是由5个7组成的，孩子们要挑出由9个4组成的那个数列才对。孩子们还要完成一项有关抑制控制的练习。当画面上出现一张小绵羊的图片时，他们要尽可能快地点击图片，但当图片是一只披着羊皮的狼时，他们就必须马上停止点击。

这些有关注意力管理和执行控制的任务，大大改善了孩子们在瑞文推理测验中矩阵问题上的表现。接受过训练的孩子的分数比未接受训练的孩子的分数高出0.4个标准差。另外，这些训练对脑电波的形状也有极其显著的影响。研究人员记录了孩子们在完成部分任务时的脑电波变化情况。负责注意力控制的那部分大脑的活动快速转变，观测到的经过训练的4岁孩子，其脑电波形状更像是6岁孩子的，而非未经训练的4岁孩子的脑电波形状。

所以我们完全有理由相信，文化使得执行控制功能比以往的时代有了更大的进步，而更加卓越的执行功能则使人们能更出色地完成与液态智力有关的任务——当然包括瑞文矩阵，可能还有其他一些测试液态智力的任务，例如韦克斯勒儿童智力测验量表中所有测试操作智力的试题。或许我们还可以期待，通过这种电脑训练任务，罗达及其同事能够帮助患有注意缺陷

多动症的儿童，注意缺陷多动症的研究人员或许还能改善未患多动症的成年人的工作记忆。

认知神经科学家阿黛尔·戴蒙德及其同事证实，日常的游戏活动也可以改善执行功能。他们为幼儿园的孩子设计了一些游戏，这些游戏可以教会孩子们如何控制注意力和约束自我，二者都是执行功能的关键组成部分。他们为孩子们制订了明确的游戏活动计划，培养其学习在利用记忆辅助设施的条件下如何按照要求快速做出反应，并且培养了孩子们迅速转换的能力。这些活动确实改善了孩子们在测试执行功能的标准测试中的表现。

瑞文推理测验的分数在55年间增加了两个标准差，韦克斯勒儿童智力测验量表中操作测试的分数增加了不止1个标准差，这是不是意味着人们的智力水平得到了极大的提高呢？可能并非如此。这些进步的确反映出人们在构成液态智力的某些方面的能力确实提高了，但这并不代表人们解决问题的能力也提高了，因为这些测试并不能对解决问题的能力进行测量。我们也不清楚这些技能究竟能在多大程度上帮助人们解决问题。

现在可以明确的是，任何宣称瑞文推理测验是一项完全与文化无关的智力测试的观点，现在都已经完全站不住脚了。对于不识字的亚马孙部落居民，或只受过一点儿西式教育的非洲人，以及生活在高度复杂化、教育高度普及化和高度信息化社会中的美国人、瑞典人或西班牙人来说，根本不能用瑞文推理测验来比较他们智商的高低——即使曾经这样比较过。

不过，这并不是说某种文化中更聪明的人在瑞文推理测验中的表现不如那些不太聪明的人。我们知道，更聪明的人不光现在在瑞文推理测验中的表现很好，在两代人之前也表现颇佳。这是因为瑞文推理测验的分数确实能够在一定程度上预测或准确预测学习能力和职业成就。只不过随着文化的变迁，例如数学教学方式的变化和电脑的广泛使用，瑞文推理测验和其他测量液态智力的测试对于每个人来说都变得越来越容易了。

随着时间的推移，被普遍认为是测量语言能力或晶态智力的两项测验的分数也得到了极大的提高。1947—2002年的55年里，韦克斯勒儿童智力测验量表类同小测验的得分的增加，相当于智商得分增加了24分。如果想要在类同小测验中得到较高的分数，你要能够回答出夏季和冬季是相似的，因为它们都属于季节——或者如果你真的很聪明，你会回答它们都有至日（夏至或冬至）。如果你回答夏季和冬季都是一年里的不同时间，只能得到一部分分数。如果你回答“夏季和冬季都是自然的一部分”或“夏季和冬季都多风”，一分也得不到。除此之外，你还要能回答出“报复”和“原谅”是相似的，因为它们都代表着你受到伤害时面临的选择。如果你回答二者都是决定，或者都是你将对别人采取的行动，只能得到一部分分数。如果你回答二者都是一种感觉或解决办法，一分也得不到。简而言之，你必须能够

从测验项目中提炼出抽象的特点，然后寻找这些特点之间最有意思且最有意义的交集。

类同小测验的分数得到了提高是否意味着在过去的两代人的时间里，我们的智商真的提高了1.6个标准差呢？当然不是，但是这种提高确实说明我们分析性思维的能力提高了，这种分析能力能够帮助我们理解并运用明喻和暗喻，同时说明了我们现在更善于运用与科学分类有关的方式给事物和事件归类。这些提高都具有非常重要的意义。

请注意，虽然类同小测验测试的是一种晶态智力，正确地回答问题也确实依赖于一个人的信息储备，但这项小测验里难度较大的试题中还是含有不少液态智力的成分。你能够记住夏季和冬季都是季节，但是你不得不当场推断“报复”与“原谅”具备哪些相同的特点，并判断哪个特点是最符合要求的。

在其他测量晶态智力的小测验中，在最近几十年里分数也得到极大提高的就是理解力小测验。在某些方面，我觉得这个小测验才最能说明人们确实变得更加聪明了。值得注意的是，现在的小孩子更能够理解为什么不用电灯时就要把它们关上，或者再大一点儿的孩子更能够理解为什么人们要纳税。理解能力的提高是惊人的——每30年一代人的时间里就增加了1/3个标准差。这样的提高应该归功于什么呢？我不清楚，但是我认为这与电视的作用密不可分。孩子们通过观看具有教育性的电视节目，例如《芝麻街》，可以学习到很多关于这个世界如何运转的知识，甚至一些纯粹娱乐性的节目也可以让孩子们学到不少东西。

理解力小测验的分数有所提高，不太可能是因为孩子们读了更多的书。为了解释这个问题，让我们看看测试晶态智力的其他三项小测验，这三项小测验的分数并没有随着时间的变化而显著提高。词汇小测验的平均得分仅仅增加了0.25个标准差。这不能说没有增加，但这与理解力小测验得分的增加幅度根本不能相提并论。不过这似乎是可以理解的，因为人们现在的阅读量比以前的人少多了。在过去20年里，不阅读任何书籍的17岁青少年所占的百分比增加了一倍。另一方面，我们从其他证据[注]确切得知，更小的孩子开始阅读的时间比35年前的同龄孩子提前了一些。从20世纪70年代初开始，美国国家教育进展评估每隔几年就对9岁、13岁以及17岁的孩子进行测试。结果表明，9岁孩子的阅读能力每一代人提高了0.25个标准差，13岁孩子略有进步，而17岁的孩子则没有任何变化。总之，美国国家教育进展评估得到的关于阅读的测试结果，与词汇小测验的测试结果是一致的。值得注意的是，考虑到最近几十年里，美国这些大孩子的课程难度大约降低了相当于两个年级的难度水平，所以虽然他们的阅读能力看似几乎没有提高，但实际上还是比预期的结果要好。

令人困惑的是，在两代人过后，信息小测验的分数实际上几乎没有任何提高。在某种程度上这一点也不足为奇，因为今天的孩子们不会花很多时间来记忆各种事实。我小的时候就被要求记住所有州的首府。不过事实证明，知道肯塔基州的首府是法兰克福而非列克星敦，对我也没有太大的实际用处。另一方面，令人感到奇怪的是，信息小测验的平均得分几乎没有任何增加，而理解小测验——对世界为什么像现在这样运转的理解——的得分却增加了不少。

韦克斯勒儿童智力测验量表中，算术小测验的平均得分在过去30年里没有增加，但其他证据显示，人们的数学技能确实提高了。这也是一个令人费解的地方。

20世纪初，只有大学和较好的高中才向学生教授几何；20世纪中叶，从十年级开始才普遍学习几何；而现在，初中后期和高中前期就已经开始学习几何，小学里也开始教授一些含有几何概念的内容或运算。在1900年时，大学四年级以前根本不讲微积分，那时美国只有不到10%的人才能上大学。到20世纪中期，较好的大学里在一、二年级或最好的公立及私立高中在高三才开始教授微积分课程。现在所有的高中在最后一年都开始教授微积分，某些精英学校甚至从初中就开设微积分这一科目了。1920年，不到20%的美国人读到高中。到了1983年，超过80%的美国人读到了高中。所以令人感到奇怪的是，在长达60多年的时间里，韦克斯勒儿童智力测验量表中算术小测验的得分居然没有丝毫提高。

不过还是有大量证据显示，学生们的数学水平确实提高了，至少是高中以前的数学成绩。美国国家教育进展评估显示，1978—2004年，9岁学生的数学成绩提高了2/3个标准差，13岁学生的成绩提高了半个标准差以上，17岁学生的成绩提高了将近1/4个标准差。韦克斯勒儿童智力测验量表算术小测验的结果，与美国国家教育进展评估得到的结果有如此大的出入，我怀疑是因为韦克斯勒儿童智力测验量表更强调按部就班地运用学习到的算术步骤，而美国国家教育进展评估则注重数学推理以及需要调动工作记忆才能解决的、需要多重操作的问题。究竟为什么两个结果存在这么大的出入，我也不敢妄下结论。

智力测试得分会无限制地一直向上增加吗？肯定不会，虽然现在还没有证据显示在美国智商得分的上升有减缓的趋势。不过在斯堪的纳维亚一带的国家，智商的增长似乎已经停止了，而在发展中国家智商可能才刚刚开始增长。在肯尼亚的某个地区，7岁儿童参加瑞文推理测验的得分在14年里提高了1.7个标准差，语言智商提高了0.5个标准差。由于参加测试的儿童只有7岁，并且刚刚开始上学4个月，如此大的提高不太可能归功于学校教育。这种提高和所接触的发达国家中日益丰富的大众文化（例如电脑游戏）也没有丝毫关系。合理的解释是，在这段时期里，家长的学历普遍显

著提高，孩子们的营养明显改善，钩虫病的发病率也有所降低。一项研究显示，在25年的时间里，多米尼克境内加勒比岛屿的居民，在瑞文推理测验中的得分提高了18分，词汇测试的得分则提高了20分。

那么现在从智商的增长中，我们可以得出哪些结论呢？

1．学校肯定能让人变得更聪明。在学校里学习到的信息和解决问题的能力会提高孩子们的智商。对于智商来说，一年的学校教育等于年龄增长两岁。

2. 随着时间的变化，人们完成用来测试智力的一些任务的能力会变得越来越强。这似乎是顺理成章的，因为越来越多的人接受教育，而且教育越来越朝那些可以提高智商分数的能力的方向倾斜。另外，大众文化的某些方面也对智力提出了越来越高的要求。

3. 很明显，某些方面智商的提高（例如在理解和类同小测验上的提高）确实表明了，人们解决日常生活中出现的问题时所需要的智力得到了提高。

4. 某些方面智商的提高是非常重要的，因为这不仅能够使学习成绩取得进步，还能提高完成涉及抽象、逻辑和现场推理的任务的能力——这些都是在工业和科学中经常运用的能力。测量这种液态智力的测试包括韦克斯勒儿童智力测验量表中的积木图案、图形拼凑、图片排列和图画补缺4项小测验，以及瑞文推理测验。

5. 虽然液态智力得到了提高，但这并不代表完成日常生活中实际推理任务的能力也得到了提高。

6. 智商提高明确了一点，测试操作智力、液态智力的测试以及瑞文推理测验并非与文化没有任何关联，而现在还有一些智力研究人员仍然宣称二者毫无关系。这种测试液态智力的任务与文化的联系，比测试晶态智力的任务要紧密得多。实际上，这些测试得分的提高凸显了一个问题，即是否存在完全与文化无关的智力测试。

7. 就像学校教育能够提高智商一样，那些社会看重的技能——既可以用于日常生活，又可以用于科学、工业及其他行业——随着时间的推移也得到了改善，这就证明人可以在某些非常实际、非常重要的方面变得更聪明。

8. 最后，我们有证据来驳斥查尔斯·莫瑞做出的两个非常悲观的论断。他曾说过，即使是再完美的教育制度，也不会提高智商值分布在低水

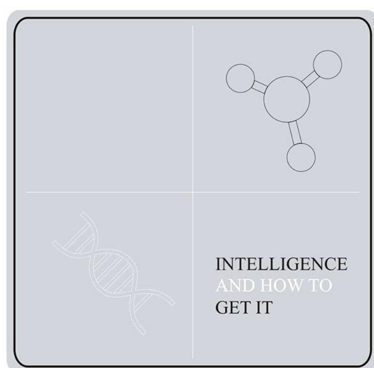
平区域的孩子的智商。但是智商值处于低水平区域的孩子的平均智商在60年里增加了1个标准差以上，并且他们在长久以来被视为智力测试黄金标准的瑞文推理测验中，得分增加了2个标准差以上。莫瑞还说过，由于智商较高的人生育的子女，数量上要少于智商较低的人生育的子女，所以整个人类的平均智力水平一定是呈下降趋势的。而证据显示实际情况恰恰相反。

在下一章中，你会读到，学校能否更好地完成让人变聪明这一任务。在第7章，你会了解到有没有可能使一个人的智商值从低水平区域提升到高水平区域。

---

1. 美国国家教育统计中心2008年的表格与图例，http://nces.ed.gov/quicktables/，2008年8月14日检索。

第4章
选择哪所学校会让孩子的成绩更优秀？



先生，为了他的教育，我可没少费心；让他从小就在街上跑，自己养活自己。这是让孩子学得机灵点儿的唯一办法呀。

——大维勒先生，选自《匹克威克外传》（狄更斯，**1836**年）

如果说在过去的20世纪里人们确实变得更聪明了，部分原因是人们受到了更多、更好的学校教育，那么能够通过改善学校教育使智力得到更大的提升吗？如果可以，需要怎么做才能让学校变得比现在更好呢？这些是摆在美国人面前亟待解决的问题，因为美国在教育成果上的表现远远落后于大部分发达国家。在常识测试中百分位排名第95位的美国学生，只相当于某些得分很高的国家中排名第75位的学生。[注]美国最优秀的学生（5%的学习大学预修课程微积分科目的顶尖学生以及1%的学习大学预修课程物理科目的顶尖学生），得分也不过相当于其他国家顶尖的10%~20%的学生。所以，即使只拿最优秀的学生来比较，也能看出美国与其他发达国家之间的巨大差距。

留给我们改进的空间还很大，所以让我们看看怎样做才能够让孩子们变得更聪明，学习成绩更优秀吧。

投入更多的教育经费管用吗？

关于花钱也无济于事这一点，倒是有一些非常惊人的证据。从20世纪70年代开始，研究人员就一直坚持认为，拨给学校的教育经费的数额与其效果并不存在密切的联系。通常采用多元回归分析的研究都会得出这样的结论。研究人员在多元回归分析中引入了大量变量，然后确定究竟是哪个变量在起作用。这些研究基本上都是要确定如果不考虑其他因素，投入的教育经费的数额是否会对教学效果产生何种影响。研究结果大致是这样：如果钱是花在东北部的学校上，这些学校的学生主要是较富裕的白人家庭的孩子，钱的数额不会对教学效果产生什么影响。如果把钱花在南部的学校上，这些学校的学生主要是贫穷的少数族裔家庭的孩子，钱的多少也不会对教学效果有什么大的影响。

很明显，只向学校投入一定数额的教育经费本身并不会对教育产生任何影响。法官有时要求社区向家境贫寒的孩子和家境富裕的孩子就读的不同学校划拨同样的经费。当经费拨给家境贫寒的孩子就读的学校后，这些额外的经费可能并没有经过认真思考和筹划就被花掉了，而且也没有证据显示学生们的成绩有什么进步。典型的例子发生在堪萨斯城。司法要求大幅增加学校的经费。学校用这些钱兴建了奥运会规模的游泳池，最先进的科学实验室，并给每名学生配一台电脑。结果是：学习成绩没有丝毫的进步。钱本身并不能提高学生的成绩，尤其是当经费的管理人能力不足或出现贪污行为的时候（就像某些大城市里发生的那样），增加教育经费就更不可能提高学生的成绩。另外，在发达国家，每名学生教育支出的多少与学生参加国际数学和科学趋势研究测试的得分高低没有必然联系，这也证明了钱本身并不能提高学习成绩。一些得分很高的国家，其人均教育支出还不到中位数，一些人均教育支出超过了中位数的国家得分却较低。

不过，这也并不意味着钱一点儿作用也没有。在后面的章节里，你会发现能够证明通过增加穷人和少数族裔人的教育经费来改善教育质量的蛛丝马迹。

教育券与特许学校有用吗？

许多教育评论家都曾强烈要求实行教育券制度——把钱发给学生家长，这样他们就可以把孩子送到私立学校去读书。但是研究人员不应该将得到并确实使用了优惠券的学生与没有得到优惠券的学生做比较，因为这里存在着自我选择的问题，情况或许是，得到并最终使用教育券的学生家长，都

是一些受教育程度较高、聪明又有进取心的父母。那些宣称教育券计划起到了很大作用的人，实际上的确是将得到并实际使用了教育券的孩子与没有得到教育券的孩子做了比较。这样的研究发现黑人与白人考试成绩的差距缩小了1/3。但实际上，这样的“收获”中至少有一部分显示的只是不同学生家长之间的差异。也就是说，有的父母很在意孩子的教育，所以最终使用了教育券，送孩子去更好的学校，而有的父母也许得到了教育券，但最终可能使用也可能不使用它。有些研究人员坚称他们可以采取一些措施控制自我选择效应，例如，比较的对象虽然仍是使用了教育券的学生和没有得到教育券的学生，但二者在很多方面都很接近。这简直就是无稽之谈。一旦为自我选择敞开了大门，怎么可能通过将比较的对象一一对应起来来规避问题呢？换句话说，一组父母接受了听起来对子女有利的帮助，而在一些方面与之相近的另一组父母可能未必会接受这样的帮助，这两组父母在其他方面也可能有很大的差异。

采用合理的方式对教育券效果进行研究后，比较得到教育券（大部分人都接受了教育券）的学生与参加教育券的抽签但未中签的学生的学习成绩，证据显示使用教育券大约可以将学习成绩提高0.1个标准差。（这种比较采取的是意向治疗的方式：对一个治疗组里面的全部学生进行分析，不管学生是否都接受了这种治疗。即使这种治疗有效，效果也因这种情况不可避免地被低估了，但是这样可以避免致命的自我选择。如果治疗组里只包括那些学生——其父母确保他们会接受治疗，将治疗组与控制组进行比较就会出现自我选择效应。）这种令人失望的结果当然不能证明私立学校不如公立学校，只是到目前为止，也没有什么令人信服的证据能够证明被研究的私立学校确实比公立学校好。

另外也没有证据能够证明，总体而言特许学校能使学生更加受益。特许学校也依靠政府的拨款，但可以不遵守政府用来管理其他公立学校的某些规定和法律。但是反过来，这样的学校必须在校章里保证取得一定的成绩。一些合理的实验随机选取了一些研究对象，然后将特许学校和其他公立学校进行比较。不幸的是，特许学校只比孩子们有可能去的公立学校略胜一筹——至少在特许学校刚刚开办的头几年里是这样。拿数学和阅读来说，特许学校中较低年级的学生在标准化考试中的成绩只前进了几名。大一点儿的学生在高年级时第一次进入特许学校后的成绩，还不如公立学校的学生。不过，也有一些证据显示，在特许学校开办了10年以后，与公立学校相比，它们或许会具备10%的优势。不过这不应该理解为特许学校不如公立学校——只是到目前为止胜过公立学校的特许学校还不多。在后面的章节里，你们就会欣喜地发现这一结论至少存在一个让人高兴的特例。

班级人数的多少有无影响？

班级人数的多少会有什么影响吗？学生人数更少的班级是不是更好呢？我们掌握的证据也是互相矛盾的。采用多元回归分析的研究人员往往会告诉我们，班级人数的多少不会影响学生的表现。而另一方面，在研究班级规模的影响方面颇有建树的研究人员，经济学家艾伦·克鲁格坚持认为，70%关于班级规模的研究都得出了肯定的结论，即班级人数的多少会影响学生的表现，而且刊登这些研究成果的杂志越重要，得到肯定结论的可能性就越大。克鲁格和我都认为，有一项研究比所有其他的研究加起来都更能说明问题。这项研究是20世纪80年代在田纳西州进行的。幼儿园和小学一至三年级的老师及学生被随机编为一般规模的班级（平均22人）或人数较少的班级（平均15人）。人数较少的班级中的学生在标准化考试中的成绩，要优于一般规模班级的学生，进步的程度根据科目和班级规模从0.19到0.28个标准差不等。换句话说，在人数较少的班级内学习，通常能使一个孩子的学习成绩的百分位排名从第50位升至将近第60位。至少在七年级前，这样的效果都非常显著。相对于中等阶层的白人孩子来说，班级人数较少对家境较差的少数族裔孩子的影响更大。

教师的水平高低有关系吗？

那教师呢？他们的水平会对学生有什么影响吗？有些教师肯定比别的教师更优秀吗？答案是肯定的。不过事实上，一个持有证书的教师未必就是一个出色的教师。即使拥有硕士学位也不能说明一个教师很会教书，这没有什么可奇怪的。

即便如此，还是有不少证据证明教师的确很重要。首先，教学经验很重要。只有一年教龄的教师和有10年教龄的教师教出来的学生，阅读成绩测验的得分平均相差0.17个标准差——也就是说教师的教学经验可以使学生成绩测验的百分位排名提高7个位次。不过请注意，通常在教师工作的第一年才会出现这些由教学经验造成的差距。所以还是要留心，尽量避免把孩子送进一个由刚走上工作岗位的教师负责的班级中去。

将某一班级中的中等学生在成绩测验中的分数，与根据前一年中等学生的分数做出的预期进行比较，高出的部分就是教师的附加值，我们可以用附加值来定义该教师的水平。如果学生的成绩超过根据前一年或前几年同年级学生的成绩做出的预期，我们可以说教师的附加值很大。按照这个定义，如果教师水平相差1个标准差，那么学生的测验成绩就会相差大约0.2个标准差。不过这只是对同一所学校内不同教师造成的差异做出的估计。既然我们知道某些学校里教师的水平平均来说要高于其他学校的教师，我们就可以肯定，这种对同一所学校内教师水平差异的估计一定是教师水平差异的最小值。将不同教师不同学校的差异纳入考虑后，经济学家埃里克·哈努谢克估计，教师水平对学习成绩的影响是0.27个标准差——这足以使

一个学生在成绩测验中的百分位排名从第50位升至第60位。教学水平等于或高于所有小学教师平均教学水平1个标准差的老师，以及教师水平等于或低于所有小学教师平均教学水平1个标准差的老师，二者教出来的学生其成绩测验分数究竟有多大的差距很难计算，不过显然会相当大。

人们往往会谈到某些教师对他们有重要的影响，这使得类似有关教师教学水平的统计数据不那么苍白无力了。很多我认识的人都认为，教过他们的老师中至少有一两位对他们产生了非常重要的影响。我们没有理由怀疑他们所说的，因为有一项关于一年级教师对学生造成的影响的研究证明，他们所讲的个人经历并不是信口开河。

A女士在一所学校里任教已经有34个年头了。这所学校主要面向社会经济地位较低的人群，其中1/3的学生都是黑人。这项研究采访了60个成年人，他们在11年的时间段里都曾经在这所学校里学习过。其中1/3的人读一年级时的教师是A女士，剩下的40名被采访者读一年级时的教师不是A女士。在这40名被采访者中，31%已经记不起一年级老师的名字了，而所有A女士教过的人都还记得她的名字。不管能不能记住教师的名字，这40人中只有1/3的人给他们老师的评价是非常好或出色。而A女士的学生中有3/4对她的评价都很高。其他老师的学生只有25%对教师用心程度的评价是A，而A女士的学生中有71%给出的评价是A。当被研究人员问及A女士是如何讲课的，我的一位同事回答：“她是带着对学生的爱在讲课。”这位同事还说，A女士表达了她对所有学生的信心，认为他们都可以学习——不会有一个学生在离开她的课堂时还学不会如何阅读。她不仅会在放学后留下来帮助那些学习吃力的学生，还经常把自己的午饭分给那些忘记带午饭的学生。

A女士的学生在小学、青少年时期和成年后与其他学生相比要出色得多。A女士教过的这些学生中，有2/3在二年级时成绩就排在了全年级的前1/3，而其他教师的学生中只有28%进入了年级的前1/3。研究还测量了这些学生成年后的状况，例如受教育程度，职业成就以及家庭状况。A女士的学生中有64%都处于最佳状况，而其他教师的学生中只有29%处于这种状态。

即使没有其他更令人信服的、量化的证据证明教师（尤其是一年级教师）的重要性，这个故事也并不能算作证据，因为这只是关于某位个别教师的特殊案例。不过确实存在其他证据。教育研究人员布丽奇特·哈姆雷和罗伯特·皮亚塔接触到了由美国国家儿童健康与人类发展研究所主持的一项规模庞大的纵向研究，即儿童早期保育研究。他们跟踪记录了900个孩子从幼儿园一直到小学一年级结束的整个过程。根据孩子母亲受教育程度所反映的社会经济状况来判断，有些孩子被认为恐怕很难适应学校的学习。另外，根据幼儿园老师报告的孩子在行为、注意力及学习成绩方面出现的问

题来判断，有的孩子可能会面临比其他孩子更多的问题。

观察人员在每间教室里都旁听了三个小时，对一年级的教育质量进行评估。所有班级按照两个相关但在统计学意义上又各自独立的方面划分等级。第一个方面是“教学支持”，是几项评估得分的总和，其中包括读写技能的教学质量，对孩子的评估性反馈的质量，师生谈话涉及教学的程度，以及对孩子责任感方面的鼓励。第二个方面是“情感支持”，也由几项评估得分组成，包括情绪气氛，对课堂的有效管理，例如拉近与学生之间的距离，以及不对学生过度干涉等。

所有班级按照教学支持的质量分为三个等级。所有的孩子都接受了一项广泛适用的能力测试——伍德科克·约翰逊认知能力测验。

把一个因母亲受教育程度较低而被视为可能会面临问题的孩子，安排到教学支持较小的班级，与被安排在教学支持较大的班级中的学生相比，一年级期末时这个孩子的测验成绩可能要低0.4个标准差。把一个具备同样风险的孩子安排在一个教学支持较大的班级，他的成绩实际上与那些父母受教育程度较高的孩子的平均成绩不相上下。

所有班级也按照情感支持的水平分为三个等级。（总的来说，教学支持级别较高的班级，情感支持的级别也较高，而教学支持级别较低的班级，情感支持的级别也较低，但也不完全如此。）有些孩子由于幼儿园时在社交和情绪方面表现不佳，被认为可能存在留级的风险。把这样的孩子安排在情感支持较小或一般的班级里，其学习成绩可能要比被安排在情感支持较大的班级中得到的成绩低0.4个标准差。根据幼儿园时的表现被认为存在留级风险的孩子，如果被安排到有较大情感支持的班级里，他的成绩和班级里那些没有这种风险的孩子并没有什么不同。

接下来要讲的这个发现或许比听上去更为重要。哈姆雷和皮亚塔在此前的一项研究中发现，在幼儿园时不善与人相处的孩子，可能整个上学期间的学习成绩都会受到影响。对一年级教师水平的研究结果表明，有类似问题的孩子如果在一年级时能够在适宜的班级里学习，他们的发展轨迹可能一开始是向下的，但可能很快就会出现逆转。

校长都清楚自己学校里教师的水平——至少清楚哪些教师最好，哪些教师最差，但是还没有证据显示校长会苦心寻访或奖励高水平的老师。实际上，对于公立学校的教师来说，这一点是很难实现的。工会的规定往往是在不同老师之间实行基本相同的工资待遇，并且待遇上的差异只能与资历、资质和学历有关。正如我之前提到的，没有证据证明资质和高学历就一定代表教学的高质量，而且除了在开始教学的第一年，资历也不能说明教师的水平。对这些情况有所了解的研究人员往往分成两个阵营：（1）

改变工会的规定，对教学水平高的教师进行奖励；（2）承认这种做法恐怕难以实现，因此只能对教师进行培训，进而提高教学水平。

大家都同意，对于改善教育，我们能做的最重要的事情之一就是提高教学质量。也许应该先从学校开始。新老师们普遍抱怨他们上的教育理论课程太多了，而经验或实习则少得可怜。另一种提高教学质量的方法可能是给优秀教师发放奖金。以色列的研究人员对两个不同的奖金计划进行了研究。一个奖金计划是根据学生的成绩向位于前1/3的学校的所有老师发放奖金。奖金数额相当于基本工资的1%～3%。另一个奖金计划是为获奖的学校提供更多资源，主要是提供教师培训或减少教学时间。这两个计划都颇有成效，使学生的成绩有所进步，退学率有所降低。但是相较而言，增加工资的奖金计划更合算。研究显示，向学校发放奖金或许能避免工会的种种麻烦。可以想见的是，如果这样的奖金被视为一种额外收入，并且发放的对象不是个人而是学校，教师和工会或许就会接受这种竞争，因为这种竞争的奖励模式是“稳赚不赔”的。不过，和评估教师的水平一样，究竟如何评估学校的质量，恐怕也是个无休无止的难题。

教学成果卓著的学校就是好学校？

直到最近，依旧没有什么令人信服的证据能证明哪些教学方法的效果更好。早就有大量关于“有效学校”的报道——这种学校出来的学生，其成绩都大大超出人们的预期。不过这样的证据并不能使人完全信服。都说好学校一定有一位好校长，这样的校长会向大家传递一种信念：相信大部分孩子都有学习的能力；会仔细地挑选教师并监督他们的行为；会想方设法将不合格的老师清理出教师队伍；重视课程设置和教学策略；关注学生的表现，以便确认学校的教学策略是否有效；鼓励家长加入到孩子的教育中来。有效学校里的教师也较容易亲近，而且通常一开口就三句话不离本行。这些教师一般更容易接受评估，并且更喜欢接受这种评估。他们会随时关注学生的表现，以确定学校的教学策略有效与否。

大部分关于有效学校的报道，都是介绍那些面向经济条件较差的人群的学校。通常认为，那些更强调基础知识的学校对于这样的人群更有效。不过另一方面，有的人也认为，有效学校的课程安排如果能像高水平的私立学校那样丰富，它们就会取得很好的效果。

总而言之，教学效果更理想的学校，一般来说都有好的校长、更有效的教学策略、更全心全意地培养孩子的教师。不过这些报道并不能让我相信是这些特点使得学校成绩斐然，反而仅仅反映出一个事实，那就是这些学校的学生更好管理。如果学生们更易管理，校长就会看起来很称职。如果学生们总是不断地惹麻烦，那么第一要务就是要抓纪律，课程安排或教师评

估只能先靠边站了。所以虽然关于好学校的故事不断上演，却没有多少无可争辩的事实能让我对如何改善表现不佳的学校抱有信心。

每项教育创新的举措都有效吗？

尽管对各种创新教育计划的投入已经高达数亿美元，对这些计划的评估也多如牛毛，但教育研究的现状令人汗颜。这些研究不仅大都不能使人信服，而且不少自封的教育计划评估者实际上并没有采取正确的实验法进行研究。实验法是指从某个人群中随机选出一些孩子，对他们采用一种教学方法，再对随机选出的另外一些孩子采用另一种教学方法，然后将两种情况加以对比。实际上，很少有教育研究遵从了这样的科学标准。

这种情况令人觉得不可思议，就好比制药公司在未经任何评估研究证实药效的情况下，就将生产的药品兜售出去，而有时这种评估研究也只是随便地让一些患有某种疾病的病人服用这种药物，然后记录下有多少比例的病人病情有所好转（但这样并不能确定有多少比例的病人病情好转根本不是由于服药）。确定一定数量的患有某种疾病的患者，然后对其中一部分使用某种药物的人进行治疗，而对另一部分人不进行治疗或使用另外一种药物治疗，然后比较两组患者的情况，这样的药物测试才算得上是合格的实验研究。但是有关教育举措的研究几乎从未达到过这一标准。谁要是说，评估教育研究不需要随机选取研究对象也可以取得一样或更好的效果，你就应该问问他们，为什么教育研究和药物测试要采用不同的标准。

不使用实验法的人给出了不少理由。有人称这是出于道义方面的考虑，接受治疗的应该是最需要接受治疗的人，而不应是随机选取的。但既然要从道义方面考虑，调查人员就必须能够证明这种治疗是有效的——可以找到一群最需要接受治疗的人，一半编为治疗组，另一半编为控制组。在这些理由中，听起来最有道理的是，实验法不能找出某种治疗中究竟哪些是产生疗效的成分。连某种治疗是否有效都不能确定，怎么可能知道这种治疗疗效如何呢？

最近，关于学校的研究至少都采用了某种形式的控制方法以进行对比。在一些研究中，调查人员说服学校同意采取某种干预措施（例如一种新型的数学电脑教学法），然后将这些学校的成绩与那些没有采取这种干预措施的学校的成绩进行比较。根据预先制定的一系列标准来衡量，例如学生所属的社会阶层和种族，这些学校的水平都很接近。这种研究虽然聊胜于无，却也仅仅只是好一点点而已，因为存在自我选择的问题：在某些研究人员并不清楚的方面，受邀采取某种干预措施的学校或许不同于那些未受邀请的学校。有自我选择的自由时，这个问题就显得尤为突出，也就是说，只有一部分受邀的学校最终采取了这种干预措施，这些学校或许在某

些相关方面优于那些虽然受到了邀请但最终没有采取行动的学校。

只比较采取某项干预措施前后的表现，也算不上是充分的研究。而有些研究在衡量某项干预措施效果大小时，比较的是采取了干预措施的学校和水平接近但未采取干预措施的学校两者的表现。我们会发现，前一种方法显示的效果一般比后一种方法要好得多。这一规律只存在一个特例，就是当采取干预措施后学校取得了非常大的进步——这种进步与没有采取干预措施的学校的表现有着天壤之别。只有在这样的情况下，人们才能够相信这样的教育举措确实是行之有效的。

“全面的学校干预措施”效果显著吗？

得到评估的教育计划中有，一部分叫作全面学校干预措施，也就是大家熟知的全面学校改革。教育心理学家杰弗里·博尔曼及其同事回顾了其中一些最有前途的计划。只要看看他们的回顾就可以对这些计划有所了解。我只介绍那些经独立第三方测试过三次以上的计划，这些测试都是比较了采取干预措施的学校与作为控制组的学校二者间的差别，与此同时，学校或学生都不是随机选取的。

一项全面学校干预计划——全体成功——在比较研究中被第三方调研人员评估了多达25次。这项计划是由一项私立基金管理的，参与的学校多达数百所。这项计划为学校准备了特殊的课程资料，包括教师行为规范手册。同时还在阅读、写作和语言艺术等科目上为教师提供了大量的培训，以及为期26天的现场职业发展训练。这项计划侧重评估学生的成绩以及学校的管理，并为每所学校配备了一名辅导人员，对在阅读上吃力的学生进行一对一的辅导，同时还强调父母的参与。最初这个项目只是面向从幼儿园到小学六年级的学生，不过现在把初中生也纳入进来了。一些学校还参与了“根与翼”计划，这个计划是对全体成功计划的扩展，包括数学、科学以及社会研究等项目。

全体成功计划似乎提高了学生们的成绩，但越独立、设计越合理的评估，其显示的效果越差——所有调研发现的平均效果值是0.08个标准差。不过，一项设计合理、完全随机选取调查对象的研究显示，所有从幼儿园到小学二年级一直参与该项目的学生的阅读成绩平均提高了0.27个标准差。不仅如此，有4项评估将“根与翼”项目也纳入进来，这些评估显示出的进步令人大为吃惊——0.77个标准差。如果想让如此大的效果值看起来可信的话，这几项研究必须都要做到独立、证据确凿，并且要完全随机选取研究对象。尽管做了如此多的研究工作，对全体成功这个计划是否有效还没有定论。只有围绕着该项目的研究能够做到证据确凿并随机选择研究对象，我们才能得出最终的结论。

学校发展计划也是一项众所周知的全面学校干预措施，是耶鲁大学的精神病学家詹姆斯·科默于30多年前创建的。该计划并未设置特别的课程或采用与众不同的教学方法，而是尽力在教职员工、家长和社区之间构建起良好的关系，并对学生采取一些健康的干预措施。该计划还成立了几个小组，负责实施针对不同学校的特殊需求制定的改革措施。不过第三方比较研究显示的效果值仅为0.11个标准差。

效果稍好一点的是针对小学的直接教学干预措施，这些小学主要面向家庭经济境况不佳的学生。这项计划的阅读和数学项目是由麦格劳·希尔出版公司负责的，同时还提供了一些教师培训。但是整个项目要与供应商签合约，由它们来提供扩展的专业课程开发和教师培训。这些课程的教学计划都写得很详细，并要求学生完成大量的写作任务。学校根据学习成绩将学生编成小组进行教学，并时常对学生的学习进度进行评估。第三方比较研究显示的效果值是0.15个标准差。

一些全面学校干预措施的成本非常高。所以我们在衡量效果大小的同时，还要考虑到性价比。为这些项目以及那些没有得到广泛评估的项目说句公道话，有些被评估的干预措施在某些学校里并没有得到很好的实施。如果不认真对待，一项干预措施就不会显示出什么效果。因为在估计效果大小时，不论执行得好的、执行得不好的，还是执行得完全走样的，都要放在一起得出平均效果值。

改善教学方法用处更大

也有一些研究是关于具体教学方法的。评估研究人员詹姆斯·库利克回顾了大量的所谓的综合学习系统。这些电脑软件系统根据学生的水平制定课程资料，记录学生的学习效果，并给出关于学习成绩的大量反馈。库利克得出了一些非常明确的结论。他发现，在关于数学项目的16项合理选定控制组的研究中，电脑化教学的平均效果值是0.4个标准差。这样大的效果对教育是非常重要的，并且除了购买电脑外这个项目的成本并不高。教写作的文字处理项目同样对提高阅读能力具有很大的影响。高年级的效果值是0.25个标准差，幼儿园和小学一年级的效果值则更高。（在提高阅读能力而不强调写作的电脑化阅读项目上，平均效果值只有0.06个标准差。）最后，电脑辅导对自然和社会科学的教学具有非常显著的成效。这些课程的平均效果值是0.59个标准差，这可以算得上是非常大的效果值了。电脑辅导还可以激发学生对所学科目的兴趣：效果值达到1.1个标准差。

到目前为止，让人印象最深刻的方法就是一种被称为“合作学习”的课堂教学方法。学生们被分成不同的小组，互相帮助学习一些材料。这种方法对任何科目都适用，只有所有学生都参与进来才算达到学习的目的，从二年

级到十二年级都可以采用这一教学方法。教育研究家罗伯特·斯莱文写了整整一本书来介绍这些方法。斯莱文对研究有一定的要求，必须有一个控制组的学生学习同样的资料，并且研究必须有适当控制，如果不能保证随机选择个人、班级和学校，就要采用对应的办法——也就是找到在很多衡量标准上都和接受这个项目的学生水平极为接近的学生。巧合的是，使用更为优越的随机选择方法和使用对应的方法得到的结果大致相同。在其中一个互助学习的项目中，即学生小组成绩分工法，每4个学生被分成一组（通常这些学生的成绩水平或种族各不相同，或者两者都不同），一起学习各种资料，然后对每个学生进行评估。斯莱文在书中提到，研究显示这种方法对标准化考试的效果超过了0.3个标准差。其中尤其令人印象深刻的是一种叫作“结构二人组”的方法：一名学生当辅导员，另一名学生接受辅导，然后双方互换角色。互助学习有很多种不同的方式，据报道，所有方式的效果都非常好。已经有足够的相关研究显示，在从二年级到十二年级的教学中学校都应该采纳这种互助学习的方法。

总结

那么我们能够有哪些选择呢？学校能够采用新方法进而更有效地让人们变得更聪明吗？可能有很多种“不能”的回答，或至少是类似“暂时还不能”或“不太可能”的回答。钱本身并不能实现多大的改变。与普通公立学校相比，教育券和特许学校也不能使学习成绩有显著的提高。教师的资格证书和文凭与学生成绩也没有什么必然的联系。不过教师的经验确实有用，至少在一定程度上如此：任教未满一年的新教师水平肯定有待提高，并且可能需要几年时间的磨炼才能学会如何教书。

教师的授课水平很关键，一些教师就是比别人教书教得好。但是现行的制度没有对最优秀的教师进行奖励，也没有淘汰最差的教师。虽然研究尚未成熟，但是至少还是有一些证据显示，对那些在提高教育成果方面取得最好成绩的学校，为学校的所有老师提供奖金就能够改善教学成果，而且很有可能有些提供奖金的计划不会涉及任何政治方面的问题。

关于“有效学校”的介绍虽然告诉了我们什么样的校长和教师能使学校的水平有所提高，但并没有说清楚其中的因果关系：乐于奉献的校长和专心教学的教师在多大程度上能够决定学校的优劣？更易教的学生在多大程度上能够让校长乐于奉献或教师专心教学？一些全面学校干预措施虽然有一些作用，但到目前为止还没有证据显示这些措施能够产生非常大的效果。

有一些非常有利的证据显示，电脑化教学很有效，尤其是在数学和科学方面的训练上。另外，使学生们朝着共同的教学目标一起努力的合作学习方式也前景广阔。

美国教育部旗下的“有效教育策略资料中心”受到了极大的欢迎。这项服务对评估研究显示为有效的干预措施进行审核。不幸的是，这些评估研究的设计通常都没有达到随机选取研究对象的实验标准，但至少比逸事报道要令人信服得多。所有申请参加有效教育策略资料中心审核的研究，至少应该是“设计合理的准实验研究”，而这一标准远远高于大多数评估研究的标准。我们可以期待，最终教育工作者们只会采纳经有效教育策略资料中心证明确实有效的干预措施，否则恐怕很难向愤怒的公众交代。

让孩子们学习解决问题的本领能提高智商吗？

如果我们倾力教授孩子们解决问题的本领，这能比让他们学习适当难度的材料更有效地提高智商吗？对于能够通过后天学习来提高智商持极度悲观看法的《钟形曲线》一书，其合著者之一理查德·赫恩斯坦在委内瑞拉进行了一项规模极为宏大的研究，为上述问题提供了非常引人注目的答案。赫恩斯坦及其同事设计了一套高难度的教材，只向七年级的学生教授一些关于解决问题的基本概念，不涉及任何具体的科目。实际上，他们想要通过给孩子们的智力工具包提供方便的工具来使孩子们变得更聪明。

孩子们学习的概念和方法都接近高中或大学水平，而不是初中水平。研究人员设计的课程每节45分钟，总共60节。这些课程的主题包括：学习分类和假设检验的基础知识，学习可以按照某些方式排序的量度的特点，探索分级归类的结构和用途；分析类比，探究简单命题的结构；理解逻辑的原则；构思和分析复杂的论点，发展能为一个问题列举出所有可能解决方案的系统化方法；学习如何权衡结果的合意性和可能性，评估数据的可信度和相关性。这些工具通常都是学习某一科目或学科的副产品，而不是老师要详述的内容。我们能不能直接教授这些工具——甚至是教给孩子们——然后说明这些工具同样适用于解决一些他们以前学习时没有遇到的问题呢？

一个字，能。正如研究人员所设计的那样，学习这些新技能极大地提高了孩子们解决问题的能力。得出的部分效果值如下：语言理解，0.62个标准差；学习如何表示“问题空间”，0.46个标准差；决策能力，0.77个标准差；创造性思维，0.5个标准差。简而言之，一般的解决问题的技巧不但是可以学习的，还可以在很短的时间内学会。

那么智力测试测量的“真正的”智力呢？学习解决问题的能力能够提高智商吗？我认为智力测试不是测量智力的唯一方法，而只是其中一种特别的方法。如果我们能够提高人们推理和决策的能力，我不在乎他们能不能在智力测试中获得更高的分数。但是实际上，在一套测试普通能力的有代表性的测试中，委内瑞拉研究中的实验组的分数比控制组高出0.35个标准差。

在一项叫作奥蒂斯—伦农学校能力测试的典型智力测试中，与控制组的学生相比，实验组的学生的智商平均提高了0.43个标准差。卡特尔文化公平智力测验与瑞文推理测验比较类似，测试试题非常具体，并且有大量测试空间能力的试题。即使是这项测试中，实验组的学生的得分也平均提高了0.11个标准差。总之，不论是用赫恩斯坦及其同事教授的那种解决问题的一般能力来测量，还是用传统的智力测试来测量智力，这一训练都收到了很大的效果。

你们可能很好奇这些调查人员后来去做什么了。他们是不是接着为八年级的学生开发了更为复杂的工具？不幸的是，委内瑞拉的政府改组了，提高中学生的智商也就不再是什么头等大事了。尽管如此，我必须说，鉴于这个项目取得的巨大成功，我很诧异也很沮丧地发现，居然没有人沿着这个项目的足迹继续走下去。

怎样做才算是有效的家教？

最后，让我们回忆一下，很多教学都采用的是一对一的家教辅导的形式。辅导者的水平有高有低，这不足为奇。实际上，马克·莱珀及其同事发现，大学生担任的家教以及其他一些对小学生进行辅导的家教，水平参差不齐，有的家教辅导几乎一点儿效果也没有，而有的则对接受辅导的学生非常有帮助。他们从这些高效家教身上找出了一些令人感兴趣的特点。

首先，一个低效的家教是什么样的呢？要成为一个低效的家教，最有效的方法就是把自己当成一个“纠错器”。明确地告诉你的学生她犯了一个错误，然后直接告诉她应该如何改正这个错误，最好使用抽象的方法来说明其中的规律。莱珀的研究中没有一个高效的家教采用这种纯认知化的纠错恣态。

怎样做才能变成一个有效的家教呢？莱珀给出了5C（Control，Challenge，Confidence，Curiosity，Contextualize）原则。

在学生身上培养一种控制感（Control），让学生感觉到自己能够掌握所学的内容。

向学生发出挑战（Challenge）——不过挑战的难度不要超出学生的能力范围。

通过大力表扬学生的成功（表达对学生的信心，使学生确信自己刚刚解决的问题难度很大）和淡化学生的失败（为学生所犯的错误找出种种理由，强调学生做对的部分）向学生灌输信心（Confidence）。

使用苏格拉底的问答法（诱导性的提问），并将问题与学生已经见过但表面上看起来却并不一样的其他问题联系起来，通过这样的方法培养学生的好奇心（Curiosity）。

将问题放在真实世界的环境里或电影电视节目中出现的环境里，使学生学会在一定情境里思考问题（Contextualize）。

专家级的家教各自都有不同的策略。他们不会改正像忘记写“＋”这样的小错误，而是当学生要犯错时，引导学生远离错误，避免错误的发生。或者有时候他们任由学生犯错，认为这对学生来说是宝贵的学习经验。他们不会因为要顾及学生的面子而降低学习内容的难度，相反，他们会换一种方式表述这些内容。大多数专家级的家教都向学生提问，问一些诱导性的问题。他们让学生来解释他们的思路。比起表现欠佳的家教，专家级的家教给出的表扬较少。据莱珀推测，低效的家教给予了学生过多的表扬，这会让学生有被评估的感觉。最后，专家级的家教永远都精心培养、体贴自己的学生。

所有有追求的家教——还有学校里那些有追求的教师都应该学习这些。

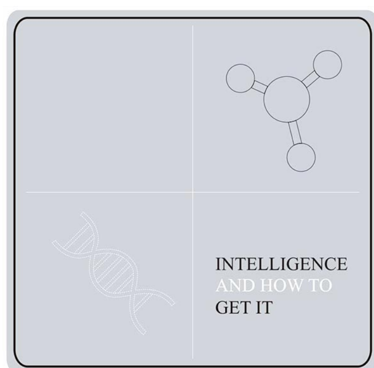
总结一下：我们知道了学校更能提高教学的质量。我们也了解了究竟哪些做法有效而哪些无效。如果学校没有参照经研究证明有效的方法来指导教学，恐怕难辞其咎。

但事实上，教学对于某些学生来说更有效果。接下来的两章里我们可以看到两类不同的学生，然后考虑一下究竟该如何改善这两章之后的章节中描述的情形。

---

1. 美国教育部1998年的数据。

第5章 家境的优劣会影响孩子的智力发展吗？



……现代社会的阶层结构基本取决于构成这些不同阶层的人的天生的智力及其他素质……

——**H. J.** 艾森克，《人类的不平等》（**1973**年）

H. J. 艾森克对于社会阶层和智力的看法代表了不少主流智力专家的观点。智商的高低决定了社会阶层的高低。穷人之所以穷就是因为他们不聪明，无论是金钱、社会阶层，还是父母的养育方式，都不能使一个人变得比别人更聪明。社会阶层主要是由出身决定的。

在第2章中，我们了解了这些观点是完完全全错误的。毫无疑问，一般来说，不同社会阶层的人拥有不同的智力基因，但是社会阶层对智力起着极大的决定作用。正如我们在第2章中看到的有关遗传度的内容，一个社会阶层较低的孩子，在中上阶层的家庭中长大与在社会阶层较低的家庭中长大，其智商平均相差12~18分。我们可以这样估计，研究中社会阶层较低的家庭位于社会经济等级最底部的10%~15%，而中上阶层的家庭则位于最顶部的10%~15%。位于社会经济等级最底部1/3的人的子女平均智商

是95分，而位于最顶部1/3的人的子女平均智商是105分。这10分的差距是由所有引起不同社会阶层智力差异的因素共同作用造成的：基因，胎儿期、围产期及产后的生物因素，以及所有与社会阶层相关的社会因素，包括居住的社区、学校及父母教养的质量。收养研究的结果表明，产后的环境因素的影响（生物因素和社会因素加在一起）可能会大于遗传因素的影响。

所以，艾森克以及其他遗传决定论者的观点简直是无稽之谈。家境的优劣会对智力产生巨大的影响。

在这一章中，我会重点说明哪些社会阶层方面的差异会直接影响智力的高低。虽然我不能将某个具体因素对智力影响的大小进行量化，但是我们知道，在这些众多的因素中，每一个因素都至少会对智力产生一定的影响。不仅如此，如果穷人能变得富裕一些，其中一部分因素就能够得到改善。我们还知道，美国的穷人及工人阶层的困窘境况存在很大的改进空间。与大多数发达国家同等阶层的人们相比，他们的经济状况差了很多。

我们也知道，不同社会阶层的父母，其养育方法也不同，他们对子女的不同养育方式自然会使孩子未来的职业及社会角色出现分化。教育家如果想帮助穷孩子提高成绩，就应该了解这些孩子究竟在哪些方面没有做好上学的准备。穷人只有改善认知功能，才能充分利用新信息经济带来的种种就业机会。

首先，我们要明确一些定义。穷人包括那些经常失业的人、长期靠救济生活的人，以及非技术工人。工人阶层包括技术或半技术工人，例如机械工人以及服务行业从业人员和办公室初级职员。我将这两个阶层统称为社会经济地位较低的人。中等阶层包括较高级别的办公室职员、教师以及初级管理监督人员。中上阶层是指专业人士和较高级别的管理人员。我将这两个阶层统称为社会经济地位较高的人。这样的定义可能有些武断，因为还有其他方法来划分社会阶层，不过这样的定义能够与在第2章中谈到的划分社会阶层的方法保持前后一致。

在这一章中，我在讨论不同社会经济地位的人的差距时，没有考虑种族因素。在下一章我会谈到种族与社会阶层是如何相互作用、相互影响的。

生物性质的环境因素

贫穷与许多生物性质的环境因素有关，这些环境因素会影响智力和学习成绩的提高。

不同社会经济地位之间存在的差异可能与营养有关，不过现有的证据表

明，母亲在怀孕期间营养不足不会影响孩子的智力，当处于饥饿中的孩子得到更好更充足的食物后，他们的智商会有所增加。西方国家不同社会阶层在营养方面的差距是否足以造成不同社会阶层在智力方面的差异，现在还不得而知，因为现在仍然有一小部分穷人处于饥饿状态中。即使不存在挨饿的情况，我们知道社会经济地位较低的孩子也极有可能会缺乏维生素和矿物质。有证据显示，缺乏维生素和矿物质的孩子在补充维生素和矿物质后智商也有所提高。

铅也会对智力造成不利影响。与生活在中等阶层家庭的孩子及郊区的孩子相比，在市中心平民区生活的孩子由于环境污染及剥落的旧油漆更易受到铅的危害。

孕妇即使摄入两盎司的酒精也会对孩子的智力产生不利影响。母亲在怀孕期间饮酒会影响孩子的注意力、记忆力和推理能力，会给孩子的日后学习造成很多的困难。和社会经济地位较高的女性相比，社会经济地位较低的女性更有可能在怀孕期间继续饮酒。

社会经济地位较低的人通常健康状况欠佳，而这会在很多方面对学习造成不利的影响。生病的孩子比健康的孩子学习起来更吃力。一些在穷人中常见的健康问题也会影响智力和学习成绩，例如牙齿健康状况不佳、受到二手烟的危害、接触污染引发的哮喘，以及视力和听力受损等。

子女出生时体重偏轻的情形在社会经济地位较低的家庭中更为常见，而体重偏轻则与低智商有关系。杀虫剂与胎儿头围较小和智商偏低有关系，虽然某些杀虫剂已经逐渐停止使用，但在社会经济地位较低的家庭仍然随处可见。

另一个我们可以量化的，或者说可以近似估计的因素就是母乳喂养。对于遗传基因最一般的孩子来说，如果采用母乳喂养，孩子的智商似乎要比不采用母乳喂养时高出6分，甚至更多。这种智商上的差异不仅出现在孩提时代，而且会一直延续至成年后。母乳中含有牛乳及配方奶粉中所没有的脂肪酸，所以母乳喂养也许会促进大脑发育。不过对于母乳喂养与智商之间是否存在因果关系尚无定论。一项研究发现，母乳喂养的孩子的智商并不比没有接受母乳喂养的兄弟姐妹的智商高。如果母乳喂养与智商之间确实存在因果关系，社会经济地位较高的人与社会经济地位较低的人的智商会因此产生约2分的差距。

另外，穷人接触到的医疗条件也较差，这不仅会导致与其他生物因素、听力和视力受损以及哮喘有关的问题进一步恶化，还会引发新的问题。社会经济地位较低的人没有医疗保险的可能性几乎是其他人的2倍。即使享有医疗保险，工人阶层的父母也不太可能带子女去医院看病，因为这样会减

少薪金收入或是因请假而受到处分。不仅如此，服务于穷人的医生也比服务于中等阶层的医生少得多，为经济状况较好的白人看病的医生的数量，是为经济状况较差的非白人看病的医生数量的3倍。

有些环境因素虽然会严重损害生物过程，但即使是在社会经济地位较低的人中间也不常见，例如铅中毒或怀孕期间摄入酒精。有些破坏程度较小的环境因素则在社会经济地位较低的人群间比较常见，例如接触受到污染的空气进而引发的哮喘。我们尚不清楚这些因素究竟会在多大程度上造成智力及学习成绩方面的差异，不过我们不能把每个因素造成的差异简单相加，然后把不同社会阶层智力差异的主要原因归结于这些因素的总和。这是因为这些不利因素彼此之间显然也存在着联系——酗酒的母亲生出的孩子更有可能接触到脱落的墙漆，牙齿健康状况也可能更差，所以我们不能把这些不利因素做简单的加和。

社会性质的环境因素

其他一些环境因素，乍看虽然不是生物性质的，却毫无疑问会对智力造成严重的破坏性影响。这些环境因素可能是通过大脑的生理机能发挥作用的。

其中一个不利因素就是，和社会经济地位较高的家庭相比，社会经济地位较低的家庭搬家的频率要高得多。这样一来，这些家庭中的孩子不得不面对经常搬家带来的压力，而且他们会转入不同的班级，这些孩子要么赶不上新的学习进度，要么不得不重复学习已经学过的知识。即使孩子是在一所学校内持续学习，也经常有其他学生转入或转走，导致班级内缺乏稳定的环境，教学效果也就难以保证。

社会经济地位较低的家庭的孩子更有可能存在行为问题，这些行为问题给周围的人造成了不同程度的麻烦。和社会经济地位较高的家庭的孩子相比，社会经济地位较低的家庭中的孩子更经常面对各种各样的不稳定因素。这些孩子所居住的社区往往更令他们压力倍增，他们生活的家庭也充斥着更多的混乱与争吵。

和社会经济地位较高的父母相比，社会经济地位较低的父母往往不会和颜悦色，不会经常鼓励和支持子女的决定，相反更有可能会严厉地指责他们的过错。发展心理学家冯尼·麦克罗伊德指出，与社会经济地位较高的父母不同，社会经济地位较低的父母教养孩子的方式往往更具有惩罚性，更令孩子紧张。

幼年时期的情感创伤会损伤前额皮质，而前额皮质与液态智力有着密不可分的联系（你可以回想一下第1章）。我们不清楚究竟多大的压力才会损

伤中枢神经系统，但很有可能有的家庭的养育方式会带来更多的压力，再加上该家庭面临的其他压力，确实会对中枢神经系统造成损伤，至少在社会经济地位极低的家庭中是这样的。而液态智力对于学习以及低年级的学习成绩来说尤为重要。

当然，不是所有社会经济地位较低的家庭的孩子都会面对这样极端的困难。实际上，大多数社会经济地位较低的家庭的孩子都生活在充满爱的家庭里，父母不仅与子女相处融洽，而且非常关心他们的身体和智力发育。很多社会经济地位较低家庭的孩子生活的社区也没有那么多危险。但即使是在最好的情况下，一般来说，社会经济地位较低家庭的孩子接触的同龄人，也不如社会经济地位较高家庭的孩子接触的同龄人那样能够激发智力的发育。不仅如此，社会经济地位较低家庭的孩子念的学校，很有可能在教师素质、班级人数、学校设施和父母的参与程度上也不能令人满意。这样看来，社会经济地位较低的家庭所处的环境会造成智商偏低、学习成绩不佳，也就不难理解了。

如何让穷人的孩子变得更聪明？

如果穷人能够变得富裕一些，智力和学习成绩方面的差距会有什么变化吗？我们知道，在美国，社会经济地位较低的家庭的孩子，其智商低于其他发达国家中同等阶层家庭的孩子。我们应该从美国各阶层贫富差距悬殊的角度，来考虑社会经济状况造成的成就差异。美国的收入差距比其他欧盟国家或日本要大得多[注]。虽然美国的人均收入比其他大部分发达国家高出25%～35%，但收入最少的1/3的工人，比欧盟国家或日本同等收入的工人要穷得多。欧盟国家中收入最少的10%的工人，其平均收入要比相对应的美国工人平均收入多44%。即使是这样的统计数字，也低估了最穷的欧洲人与最穷的美国人之间的差距。欧洲人都享有国家健康保险和其他福利措施，而美国大多数低收入者都是自掏腰包，要是负担不起就干脆不上保险。

美国的收入差距不断拉大，其速度远远超过其他任何一个发达国家。1979年，收入最高的10%的工人，其时薪是收入最低的10%的工人时薪的3.5倍。26年后，前者的时薪是后者的5.8倍。对于有子女的家庭，在1979—2002年，收入最低的1/5的家庭，其税后收入仅增加了2.3%。相比之下，有子女的中等阶层家庭的税后收入，在同一时期里增加了17%。联邦政府规定的最低工资在1997—2006年没有任何增长。虽然最近已经通过了一项大幅提高最低工资的法案，但如果按照美元的实际价值计算，2009年的最低工资即使得到最大幅度的提高，也仅相当于1968年最低工资的73%。

如果按照经济合作与发展组织收集的在读写技能、数学和科学方面的得分

来衡量，美国不仅在收入方面存在巨大差距，而且不同社会阶层之间存在的技术差距也比很多欧洲发达国家大得多。美国社会经济地位最高的1/4的人的得分，比最低的1/4的人的得分几乎高1个标准差。而斯堪的纳维亚国家在这一比较中的结果是不到2/3个标准差。造成这一差距的主要原因是，斯堪的纳维亚国家中社会经济地位最低的1/4的人的得分要高一些。美国社会经济地位较低的人与社会经济地位较高的人在阅读与数学技能方面的差距，比研究中的其他22个工业化国家都要大。美国与韩国的差距就更明显了：在韩国，社会经济地位最高的1/4的人与社会经济地位最低的1/4的人，平均学习成绩只相差1/3～1/2个标准差。

实际上，在美国，成绩最低的1/4的人与最高的1/4的人之间的差距，更接近发展中国家的差距水平，而非发达国家。

所以我们完全有理由相信，如果社会经济地位较低的人收入有所增加，美国人在智商和收入方面的差距就有可能缩小。低收入会引发很多问题，不仅会影响营养与健康状况，还会因频繁搬家而扰乱正常的生活和学习，更会降低对教育的期望。这样就陷入了恶性循环，收入低造成学习成绩差，学习成绩差造成这些孩子未来的社会经济地位低，社会经济地位低造成他们在劳动力市场上的价值低，这就使他们的社会经济状况进一步恶化。

总之，如果我们想让穷人变得更聪明，我们就需要先设法使他们富起来。

父母的养育方式不同也会造成孩子智力水平的差距

但其他造成测试得分产生巨大差距的因素就不是光靠钱能够解决的了。其中一些因素涉及社会经济地位较低的父母养育子女的方式。他们的方式不仅使孩子们很难从家庭内部学到任何知识，而且还给孩子们在学校里的学习造成了更多的困难——简而言之，就是认知文化的问题。

社会经济地位较高的父母从子女很小的时候就开始未雨绸缪。当孩子还躺在摇篮里时，这些父母就开始着力培养孩子善于提问、善于分析的能力，而这些都是成为一名专业人士或高级管理人员所必不可少的素质。而社会经济地位较低的家庭的养育方式，不是在培养未来的医生或首席执行官，而是在塑造未来的工人。只要他们长大以后听话守规矩，让雇主满意，别人也不会对他们进行进一步的评估或抱有什么其他指望。

堪萨斯大学的心理学家贝蒂·哈特和托德·莱斯利耗费了大量的人力、物力和财力，进行了一项关于儿童言语行为差异的研究。研究对象包括白人精英阶层的子女，白人工人阶层的子女和黑人工人阶层的子女，以及社会底层靠救济维持生活的黑人的子女。研究人员在他们的家里对孩子及其父母进行了数小时的观察。在这一章，我会把重点放在精英阶层与工人阶层的

家庭之间的差别上。

身为专业人士的父母会比一些工人阶层父母花更多的时间和子女谈话。母亲会对孩子说很多话，不断告诉孩子她对这个世界的看法、她自己的经历和情绪，并询问孩子的需要和爱好。一些工人阶层父母则和孩子谈话较少，即使开口也大都是命令，而命令是不会激发孩子们智力方面的好奇心的。精英阶层的家庭会让孩子加入饭桌上的讨论中，并鼓励孩子思考所讨论的问题，同时注意扩大孩子的词汇量。相比之下，一些工人阶层父母只顾自己聊自己的，从来没有想过孩子也可能对讨论的话题感兴趣，或也想要发表一下自己的看法。

身为专业人士的父母每小时大约对子女说2 000个单词，而一些工人阶层父母则大约说1 300个单词。到三岁时，专业人士的子女已经听过3 000万个单词，而一些工人阶层的子女只听过2 000万个单词。由此产生的词汇方面的差距非常明显。到三岁时，专业人士的子女掌握的词汇量大约是一些工人阶层子女的1.5倍。

不同阶层的家长在情感方面对待子女的方式也不同，而这些方式很可能会影响到孩子智力兴趣的培养和智力水平的提高。身为专业人士的父母每责备孩子一次，接下来就会有6次鼓励，而一些工人阶层父母每责备孩子一次，只有2次鼓励。父母对孩子鼓励的多少与智力发展和孩子信心的树立是息息相关的——身为专业人士的父母在这方面的做法是无人能及的。

中等阶层父母的养育方式：主动分析和求知

我们所了解的关于社会阶层及孩子学习读写技能与对学校生活的准备之间的关系，大部分都来自人类学家雪莉·布莱斯·希斯关于社会化的经典研究。希斯在北卡罗来纳州的一个城镇里花费了数月的时间，研究中等阶层的白人家庭（父母中有一方是教师），工人阶层的白人家庭（大多数家庭中的父亲在当地的一家纺织厂工作），以及工人阶层的黑人家庭（主要是农业工人、工厂工人或者靠救济生活的人）。希斯实际上就住在这些家庭里，从早到晚进行观察，并跟踪记录孩子们在学校的情况。她发现，无论是在与读写技能相关的活动中，还是在准备上小学的充分程度上，这三组孩子之间都存在着巨大的差异。尽管希斯开展研究的时间是在20世纪70年代末期，研究对象也只是某个社区中的一小部分家庭，但是后来进行的参与者更多、范围更加广泛的大量研究都显示，不同社会阶层父母在养育子女的方式上的差异，与希斯研究的结果几乎完全一致。接下来，我会主要介绍一下希斯的研究以及安妮特·拉赫的研究。

与一些工人阶层父母相比，中等阶层的父母会花更多的时间读书给子女听。中等阶层的家庭中有很多儿童读物。孩子只有6个月大的时候，父母

就开始念书给他们听，而那时孩子刚刚能够坐着。中等阶层的父母读书给孩子听，不仅仅是给孩子提供一种娱乐方式，更是要鼓励孩子把在书上看到的内容与外部世界的情况联系起来。父母有意把从书上读到的与日常生活和周围的事物联系起来。（“比利有一只黑色的小狗。你认识的人里还有谁有一只黑色的小狗呢？”“这是一只知更鸟。我们在哪里读到过知更鸟呢？它们都吃什么呢？”）家长们还鼓励孩子分析读到的内容。（“接下来会发生什么呢？她要做什么呢？为什么她要这么做呢？”）

从孩子很小的时候开始，中等阶层家庭的父母就常常会根据读到的内容发问，孩子们也知道应该怎样回答这些问题。父母让孩子们找出物体的特点，并且教他们如何根据这些特点对物体进行分类。（有一次我在飞机上正好坐在一对父子后面，孩子只有三岁大。父亲拿着一本图画书，正在问儿子哪些物体是长的，哪些物体是短的。“不，贾森，睡衣是长的。”）中等阶层的父母还问一些“什么”式的问题（“那是什么”，“博比要做什么”），接着问一些“为什么”式的问题（“为什么博比要那样做”），然后问一些要求孩子做出评估的问题（“你更喜欢哪个士兵”，“为什么你更喜欢他”）。大人们鼓励孩子谈论书上的内容，甚至让他们根据从读到的故事中受到的启发来编一些新故事。

中等阶层的孩子已经完全做好了上学的准备。他们知道如何从书本上获取信息，并且期待从中得到乐趣，而且他们也熟悉如何回答所谓的已知答案问题，即提问的人已经知道答案的问题。低年级的课程对这样的孩子来说很轻松，并且他们对要求分析和评估能力的高年级学习也准备得更加充分。

工人阶层父母的养育方式：被动接受

工人阶层的家庭里也有一些儿童读物——《金色童书》或《圣经》的一些故事，不过也许总共就十来本书。家里的墙上装饰着几幅画，上面有几首儿歌，也许还挂着一串风铃。家庭成员、朋友、邻居们都会和孩子说上一两句话。

虽然一些工人阶层的父母也念书给孩子听，也会就书的内容提问，但是没有有意识地把书中的内容和外部世界联系在一起。如果书上画着一只小鸭子，母亲或许会问孩子是否还记得有一次他在湖中看见的鸭子，不过她不会接着解释书上画的毛茸茸的小鸭子与湖中完全长大的野鸭有什么联系。大约在三岁以后，父母就不再鼓励孩子们就读到的内容进行一问一答。相反，孩子们听到的是“现在你要用心听……”孩子们要做的就是专心听父母读书，如果孩子们想要谈谈自己的感想或是提上一两个问题，都会被家长视为是一种打断。

（费城的一项研究揭示了不同社会阶层读写技能差距的外在表现和内在原因。在那些几乎所有成年人都受过大学教育的地区，书商为每100名儿童准备了最多1 300种不同的儿童读物；而在爱尔兰和东欧的蓝领工人居住的社区里，每100名儿童最多只能买到30种不同的儿童读物。估计再也没有什么能比这组数字更触目惊心地反映出不同社会阶层在读写技能上的差距了。）

中等阶层家庭中的各种活动都是用言语来引导的。一位中等阶层家庭中的父亲，如果要给孩子示范如何用球棒击打棒球，他会说："用你的手指交叉握住球棒底部，保持大拇指放在这个位置；握球棒的手不要超过这条线；别把球棒搭在肩膀上，要与肩膀上方保持两三英寸的距离。"而一些工人阶层的孩子就无法获得如此详细的指导，也没有将口头指导转化为身体动作的经验。相反，他们的父母只是简单地告诉他们"这样做"，"不对"，"应该像这样"。当要开始一项新游戏时，中等阶层的家庭都会大声读出游戏规则，并表明自己的看法。而一些工人阶层的家庭往往只是胡乱猜测游戏的规则，然后就开始游戏，一边玩一边制定规则。中等阶层的母亲都会按照菜谱做菜，她们或许会大声把菜谱读出来，这样孩子就能把听到的和实际使用的原料及正在进行的步骤一一对应起来。而一些工人阶层母亲则不怎么按照菜谱做菜，即使使用菜谱，她也不太可能为孩子创造机会，让他们把菜谱和手边的原料联系起来。

工人阶层的孩子进入学校之前也做了足够的准备，这些准备让他们在低年级时的表现还说得过去。他们通常都认识字母表；能说出颜色的名称，会数数；能告诉别人家庭住址和父母的姓名；能够安安静静地坐着听故事，也知道如何回答"什么"式的与故事有关的问题，但当被问到"你喜欢这个故事的哪一点呢"，就没有几个孩子能马上回答出来。当被问到"换成你，你会怎么做"时，通常他们都被难倒了。随着年级越来越高，学习的重点变成归类、分析和评估时，这样的孩子就会处于明显的劣势。当要求他们写一个故事时，他们往往只是重复父母念给他们听的某个故事。当被问到实际没有发生但在其他条件下可能已发生的事情时——"如果比利没有告诉警察发生了什么事，会发生什么呢"——他们就完全摸不着头脑了。

遇到这些困难的孩子到了初中很容易丧失信心，并产生厌学情绪，一步步走向辍学的边缘，最后无法顺利从学校毕业。

希斯的研究发现，不同社会阶层在培养孩子读写技能和为上学做准备的方面存在差别，这一发现可以帮助我们理解孩子们的智力和学习技能在暑假期间发生了哪些变化。中等收入家庭的孩子的智力及各种技能在这段时间通常没有进步也没有退步。但是社会经济地位较低的家庭中的孩子，各项技能都有所退步，因为不能指望他们的家庭在暑假这段时间里能够提供中等阶层家庭所给予的一样多的文化激励。中等阶层家庭的孩子在暑假中不

会退步，因为他们肯定会参与一些更能使人受到教育的活动，例如阅读或听父母念书，在餐桌上倾听具有启发性的谈话，参观博物馆和动物园，参加一些艺术、音乐甚至学习方面的课程。一项研究发现，处于从学前班到小学一年级的过渡期的孩子中，社会经济地位最高的1/5的家庭中的孩子，实际上在暑假期间有一定进步，而社会经济地位最低的1/5的家庭中的孩子，则有明显的退步。所以，中上阶层家庭的孩子和较低阶层家庭的孩子在智力和学习技能方面存在差距，主要要归因于较低阶层家庭的孩子在暑假里累积起来的退步。这些孩子在接下来的学年里也难以一下子追赶上来。

遗传决定论者读到这里或许会想：“我们怎么知道这些社会化方面的差距实际上会决定孩子的智力和学习成绩呢？我们怎么知道这不是仅仅因为社会经济地位较高的人的孩子比社会经济地位较低的人的孩子更聪明呢？我们怎么知道这是因为环境的影响，而不是因为他们幸运地继承了父母的优秀基因呢？况且拥有更聪明遗传基因的父母确实会做出刺激孩子智力发展的种种努力，他们的基因使他们乐于做这样的事情，并且这样做能够在孩子们身上收到更好的效果，因为孩子们变得更聪明了。”

毫无疑问，遗传决定论者所说的确实可以用来解释目前发生的部分实际情况。在一定程度上，环境的差异是因为社会经济地位较高的人拥有高智商的基因，而社会经济地位较低的人往往拥有低智商的基因。

不过想想看，遗传不可能是造成智力和成绩差距的主要原因。纯粹由环境因素造成的社会经济地位较高和较低的群体（有可能大体上就是最高的15%和最低的15%）之间的智商差距是12～18分，这并没有给基因留下太大的空间。社会经济地位最高的1/3的人与最低的1/3的人之间的差距是10分，而我们知道其中一大部分是由于环境差异造成的。

在后面的章节里我还会讨论，我们究竟能在多大程度上提高工人阶层和较低社会阶层家庭的孩子的智力，所以我们一定要把环境能够对智力起到很大的决定作用这一点牢记于心。根据我们前面已谈过的收养研究结果，以及我在本章中详细列举的种种环境因素，我们知道还有很大的空间留给我们，即通过改善环境来做一些改变。

改善穷人的经济状况一定会使情况有很大改观。不过，即使经济状况能够得到改善，在经济状况好转后的第一代人里，来自社会经济地位较低家庭的孩子，其智力和学习成绩也不一定会有很大提高。将军们总是要为最后一战做准备，而家长们对子女进行社会化教育时，参考的都是他们父母的生活状况，而不是他们自己或子女的生活状况。事实上，一些证据显示，在家庭收入有所增加的第一代人里，孩子智力水平提高的程度极其有限。单纯由收入改善引起的智力提高，都是通过一代又一代人渐进地实现的。

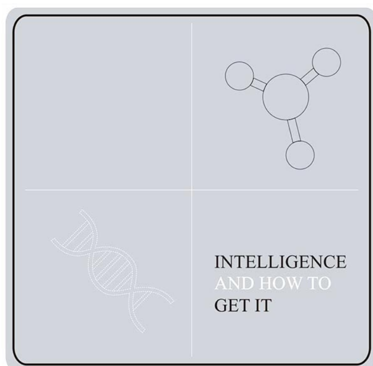
幸运的是，正如你将要看到的那样，学校可以加快缩小社会阶层之间智力差距的速度。如此一来，社会经济地位较低的家庭的孩子，就能够更好地把握信息时代经济带来的种种机遇。

不过在此之前，还是让我们先来看看不同种族的人在智力和学习成绩上的差距。造成这些差距的原因中，有一部分与造成社会经济地位差异的原因相同，其余的则不尽相同。

---

1. 这一段和下一段中出现的经济及读写能力的统计数字，来自理查德·弗里曼出版的一本关于美国劳动力市场的新书。

第6章
黑人的智力天生就不如白人吗？



讨论有关种族与智力的话题一直是一种禁忌，所以我们从未就这个话题展开过自由的讨论。这恰恰使得现在大部分测试专家坚信种族间的智力差异是由基因造成的。某种观点如果不能得到公开的讨论，就不能被驳倒。

——托马斯·索威尔（**1994**年）

黑人孩子似乎……在潜意识里认为黑人不如白人。我觉得我会对他们造成不利的影响。

——人类学家约翰·奥格布采访的一名俄亥俄州高中黑人学生（**2003**年）

黑人与白人的智力是否天生就存在差距的问题，可以追溯到1 000多年前。那时正值摩尔人注入侵欧洲，摩尔人猜测欧洲人恐怕天生就不能进行抽象思维！（1 000多年前，欧洲南部的居民对北部的居民心存种种疑问。西塞罗警告罗马人，不要把英国人买来做奴隶，因为训练他们太困难了，不过恺撒认为他们“至少还能做点儿粗活”。）但是到了19世纪，可能大多数欧洲人都认为他们的智力技能天生就胜过非洲人。

20世纪初发展起来的智力测试则强化了这种遗传论的观点。由于白人在测试中的分数高于黑人，不少心理学家在智力主要来自遗传的假设下，认为黑人与白人之间智力差距的根源在于遗传基因。

几十年来，白人在智力测试中的平均成绩是100分，而黑人则是85分——差距是15分，也就是整整1个标准差的差异。如果这一差距的主要原因或全部原因在于基因，那么这会对美国社会造成极其可怕的影响。即使在公平竞争的环境下，仍会有比例远远高于白人的黑人难以养活自己，而且能够在商业或专业领域中取得成功的黑人的比例也要比白人少得多。

在这一章中，我会针对用来证明黑人与白人的智力差异是由基因所致的证据进行讨论。同时，我还会向大家说明，现在社会上仍然存在不少巨大的社会障碍，以及一些典型的美国黑人的社会风俗，这些都严重地阻碍了黑人在教育和职业方面的进步。

黑人和白人的智力差距不是基因造成的

关于种族间的智力差距究竟是部分由基因引起的还是纯粹由环境造成的，外行们的看法迥然不同——行为科学家们也是如此。我认识的一些外行以及科学家认为，根本不用证明就能够知道，种族间的差距不可能是基因造成的，但是这样的论断没有任何根据。有可能有100种方式导致不同种族的智力天生就存在差异——不是黑人智商更高，就是白人智商更高。这是个要靠实验证明的问题，而不能因为相信所有种族都是平等的，就认为不同种族之间不存在天生的智力差异。实际上，在这个问题上有大量通过实验取得的证据。

1994年，心理学家理查德·赫恩斯坦和政治科学家查尔斯·莫瑞出版了《钟形曲线》一书，认为黑人的智力水平明显低于白人。他们自称以不偏不倚的态度看待所得到的证据，而这些证据指向的结论就是，种族间的智力差异主要是由基因造成的。查尔斯·莫瑞（赫恩斯坦几乎在《钟形曲线》一书出版的同时就去世了）不断强调，他们并不确定种族之间的智力差距究竟有多少是由基因造成的，但书中出现的证据明显向遗传决定论的观点倾斜。黑人与白人的智力天生就存在差距，公众和科学界的大多数人都认为这就是该书支持的观点。

赫恩斯坦和莫瑞在《钟形曲线》一书中给出了支持这一观点的论据。J. 菲利普·拉什顿和阿瑟·詹森在2005年出版了一份长达60页的报告，对赫恩斯坦和莫瑞给出的证据进行了核查，也提出了支持这一观点的相关论据。在这一章及附录二中，我会将这些论据一一呈现，不过这个附录不是为一般读者准备的，可能更适合那些想要看看我是如何驳斥遗传决定论观点的专家。

黑人与白人在智力上的差距并不是一些显而易见的人为原因造成的，例如黑人不熟悉较为正式的英语，或者他们缺乏参与智力测试的积极性，又或者他们的老师或智力测试人员对于他们的表现没有抱太高的期望。

不过有足够的证据显示，有时候黑人在智力测试和成绩测验中的表现不佳，往往是因为这些场合凸显了他们的种族，并带来一种“刻板印象”，导致他们无法好好表现。如果在更放松的环境里，也就是不需要担心白人测试人员对他们存有刻板印象时，他们的表现会更好。社会心理学家克劳德·斯蒂尔和乔舒亚·阿伦森证实了这一点，随后无数的研究也都证明了这一点。当黑人与白人一起参加测试时，并且测试人员明确表示测试的是智力时，黑人极有可能无法发挥出正常水平。如果测试的题目是拼图游戏，并且黑人确信在这项测试中他们和白人的表现会一样好时，黑人的表现通常会胜过在承受更大压力的环境下的表现，有时这两种环境下的表现还会有很大的差别。

不过实际情况是，一般而言，黑人在学校或工作上的实际表现并不如他们在智力测试中的分数所预测的水平。至少直到1980年詹森调查这个问题时，黑人的学习成绩和职业成就实际上低于根据他们的智力测试分数预测的水平。处在同一智力水平上时，白人的学习成绩和职业成就都高于黑人。

大体来说，黑人的社会经济地位较低，而社会经济地位较低的人智力测试的得分也不会高。但这个事实本身并不能说明遗传度的问题，因为尚不清楚社会经济地位较低在多大程度上会导致智商较低，或是智商较低在多大程度上会导致社会经济地位较低。不过我们可以确定的是，在社会经济地位的各个等级里，黑人的智商都比同级的白人低，所以社会经济地位不可能是造成两者智力上存在差距的全部原因。

遗传决定论者提供了大量证据证明黑人与白人在智力上的差距是由基因决定的，但是所有这些证据都是间接证据。例如，有证据显示脑容积的大小与智商的高低有关，大脑容积与智力的关联度或许高达0.4。而且根据拉什顿的大量研究，黑人的脑容积小于白人。

不过，脑容积与智商的关联度并不能表明两者之间存在因果关系。如果脑容积大就意味着更聪明的话，那么我们应该能够在家庭成员之间找到这种关联性。因为遗传而脑容积更大的兄弟姐妹，在智力测试中的分数应该更高，但实际上这种现象并不存在。

不仅如此，根据拉什顿和詹森的报道，男女脑容积的差异比黑人与白人脑容积的差异大得多，但是男女在智力测试中的平均得分几乎没有什么差别。在厄瓜多尔的一个地区里，遗传基因导致一群人头部畸形，他们的头

都非常小——因此脑容积也很小——但是他们的智力水平却和他们未发生畸形的亲戚们一样高，而且学习成绩远远胜过他们居住区里的大多数人。过去几千年里人类进化的趋势恰恰也是脑容积逐渐缩小。完全是出于兴趣，我注意到爱因斯坦的大脑只有1 230克，明显小于拉什顿的研究中黑人的平均脑容积。

就像关于大脑容积这样的证据一样，遗传决定论者们提供的证据大多是间接证据。如果我们掌握了关于智力差距的更直接的证据，就没有必要依赖于这些间接的研究结果。我们可以借助自然实验来证明一下，和非洲的基因相比，来自欧洲的基因是否能使一个人具有更高的智商。美国黑人的基因中大约有20%来自欧洲，也就是说，任何黑人的基因都处于100%纯非洲基因和欧洲基因占大多数的范围内。如果欧洲的智力基因更优秀的话，那么拥有更多欧洲基因的黑人的智商，就应该高于拥有更多非洲基因的黑人。

对这一假设进行证明的一个测试，就是确定带有典型欧洲人遗传体貌特征的黑人智商是否更高。结果是，浅肤色和带有典型高加索人特征的五官——都是用来衡量一个黑人拥有欧洲血统多少的指标——与智力的联系非常微弱（相关度介于0.10～0.15）。即便如此，我们应该预见到两者的联系很可能会更大一些，因为这些特征使这样的黑人在社会上具备一定的优势。

另一个关于遗传假设的测试来自第二次世界大战。那时美国的白人士兵和黑人士兵都与一些德国女性育有子女。因此，有些孩子的基因100%都是欧洲人的基因，有些孩子的基因则主要是非洲人的基因。在童年即将结束时，对这些孩子进行智力测试，父亲为白人士兵的德国孩子，平均智商是97分；父亲为黑人士兵的孩子，平均智商是96.5分——几乎没有什么差别。

如果欧洲人的基因具备优势，那么最聪明的黑人孩子的基因就应该大部分都是欧洲基因。但是一组调查人员将芝加哥所有学校里最聪明的黑人孩子找出来，并询问他们父母和祖父母的种族后发现，这些孩子的基因中含有欧洲基因的比例并不比黑人的整体水平高。

血型检测一直被用来评估黑人个体拥有欧洲基因的多少。血型化验并没有显示含有欧洲基因的多少与智商高低之间有什么联系。同样，在黑人中与高智商联系最密切的血型也并不是欧洲人中最常见的血型。

对“遗传vs.环境”问题进行测试的一个方法就是研究在白人的环境中长大的黑人孩子。如果黑人在智力上的劣势地位完全是由环境所致，那么在白人环境中长大的黑人孩子的智商应该高于在黑人环境中长大的黑人孩子。在

这一点上，遗传决定论者引用了20世纪80年代的一项研究。这项研究显示，由白人父母收养的黑人孩子的智商低于白人父母收养的白人孩子。被白人父母收养的不同种族的混血孩子的智商，介于黑人孩子与白人孩子之间。但是，正如研究人员承认的那样，这项研究有很多缺陷，例如，黑人孩子被收养时的年龄远远大于混血孩子被收养时的年龄，而收养时年龄的大小与智商的高低之间是有一定联系的。

发展心理学家埃尔西·穆尔进行了一项更大范围的收养研究。这项研究将黑人或白人中等阶层家庭收养的黑人孩子和混血孩子进行了比较，并没有发现这些孩子的智力有什么差别。

一些最近的重要研究帮助我们找出了究竟是哪些因素造成了不同种族在智力上的差距。心理学家约瑟夫·费根和辛西娅·霍兰对社区学院里的黑人和白人学生进行了测试，评估这些学生了解多少词汇和概念，以及运用这些词汇和概念学习和推理的能力。白人学生掌握的词汇和概念比黑人学生掌握的要多得多，但当测试参与者学习新单词的能力时，不论是学习字典上的定义还是在文中出现该词的含义，黑人学生的表现并不比白人学生逊色。

白人学生对格言的理解能力更强，识别事物相似性的能力也更强。尤其是当需要使用一些词汇和概念才能找出相似之处时，白人学生明显更胜一筹，而黑人学生往往根本不了解这些词汇和概念。不过当使用双方都了解的单词和概念测试这种推理能力时，双方的表现并没有什么差别。在每个种族内部，先验知识能够体现学习和推理能力。但是在不同种族之间，先验知识的不同并不能代表学习或推理能力的高低。

如果不同种族在学习和推理能力上没有差距，那么知识上的差距就不太可能是由基因造成的。更为可能的解释是，知识上的差距完全是由环境造成的。（尽管如此，我一定不会争辩说知识上的差距不属于智力上的差距。因为智力水平的高低在很大程度上取决于了解词汇和概念的多少。）

费根和霍兰的研究极为重要，但是还需要使用各种不同的材料和不同的参与者反复进行测试，我们才能完完全全地相信由此得出的结论。

关于智力差距是否由环境所致的问题，最具说服力的证据来自弗林关于最近几代人智商变化的研究。这项研究，也就是我们在第3章里介绍过的，证实了1947—2002年所有发达国家的人口的智商有了显著提高。在美国，智商提高了18分。基因不可能在如此短的时间内使智商发生这么大的变化，这一定是强大的社会因素带来的改变。如果这样的社会因素使整体人口的智力水平随着时间的推移而发生变化，它们也一定能在特定时间里使其中某个族群发生很大的改变。事实上，黑人现在的智商就高于1950年

时白人的智商。如果黑人的智力基因不如白人的智力基因，这样的情况就不可能发生了——除非你说黑人今天所处的环境比白人在1950年所处的环境更有利于培养出较高的智商。恐怕不少人都要做这样的争辩。

最后，既然我们有充足的理由相信黑人所处的环境改善的速度远远快于白人的，那么如今黑人与白人的差距和过去相比应该缩小了才对。而实际上，在过去的30年里——在很多方面，这30年比此前所有的时代都更有利于黑人的发展——12岁黑人儿童和白人儿童的智商差距确实已经由15分缩小到了9.5分。美国教育进展评估长期趋势测试也显示出黑人取得了同样的进步。白人在阅读和数学方面的提高并不明显，而黑人却取得了惊人的进步。在美国教育进展评估的测试中，黑人与白人的差距缩小的程度大致上等同于迪肯斯和弗林在智力测试中发现的5.5分。

黑人与白人的智商差距缩小了这么多，其重要性我们怎么估计都不为过。当人口的平均智商发生改变时，钟形曲线两端的差异就会很大。如果白人的平均智商比黑人的平均智商高出15分——这是两者过去的差距——那么白人智商超过130分的概率是黑人的18倍。这就意味着，那些成功的医生、科学家和专业人士在白人中所占比例要远远多于黑人。但如果差距是10分，白人和黑人的这一比例就接近6：1。虽然取得最高成就的还是白人占多数，但是黑人与白人之间的差距就不会如此巨大。智力测量表另一端的情况也是如此。如果黑人与白人的智商差距是15分（黑人低于白人），那些无法养活自己的黑人的比例要比白人高得多；而如果差距是10分，黑人的比例只是可能比白人更高一些而已。

阻碍黑人取得成就的因素在于环境

那么为什么多数黑人在智力测试中的分数一直很低，而且学习成绩不佳，事业上也没有取得什么成就呢？证据显示，基因在这中间并没有起到什么作用。所以我们可以认为这全部是环境因素所致。

首先，那些影响社会经济地位较低的人的能力和成就的问题，在黑人身上更加严重，穷人中黑人所占的比例远远超出他们在整个人口中的比例。我来提醒一下这些潜在的问题都是什么：胎儿期护理及营养不足，不采取母乳喂养，饥饿，维生素和矿物质不足，铅中毒，胎儿酒精中毒，医疗条件差，更易接触到污染进而导致哮喘，情绪创伤，学校和居住的社区环境较差，住在同一社区的同龄人不出意料地也不太令人满意，频繁搬家以及因此造成的对教育的不良影响。和白人穷人相比，这些问题对于处于社会底层的黑人来说更为严重。

除此之外，还有一些问题也导致了黑人境况的进一步恶化。2002年黑人家庭的收入是白人家庭的67%，但是黑人家庭的财产仅仅是白人家庭的

12%！造成这种差距的部分原因是黑人很难得到贷款——白人可以利用财产创造更高的投资回报，而黑人则被排挤在外。所以大体上，在就业不足或失业的情况下，黑人几乎没有什么财产可以用来暂缓经济上的压力。社会阶层较低的黑人几乎一无所有。另外，这种财富的差距也反映了一个事实：位于中等阶层的美国黑人也不过是刚刚迈入这一阶层而已。因此我们可以推测，他们养育子女的方式不同于典型的中等阶层白人家庭，而更接近于社会经济地位较低的社会阶层。

美国女性中黑人未婚生子的比例是72%，而白人的这一比例是24%。这一统计数字意味着黑人孩子会遇到更多的问题，特别是单亲家庭的贫困率要远远高于双亲家庭。或许同样值得注意的是，这种家庭只有一个成年人，而家里的成年人越少，这个家庭的环境就越不利于孩子的智力发展。

不过黑人的经济状况也并非一成不变。在美国黑人中正在形成两个不同的发展趋势，一个有利，一个非常不利。

有利的趋势是，越来越多的黑人迈入了中等阶层，早已成为中等阶层的黑人则正在巩固他们的经济地位。机会均等行动计划很可能在其中起到了一定的作用（不过有些研究人员认为，黑人早在机会均等行动计划普及之前就已经取得了重大的进步）。

非常不利的趋势是，处于较低阶层的黑人的经济状况进一步恶化，不光是黑人，所有处于社会经济地位较低阶层的人都面临着窘迫的境况。回想一下我们在上一章中谈到的，和过去相比，穷人和工人阶层的实际收入反而降低了。机会均等行动计划对社会经济地位较低的黑人几乎没有起到任何作用。所以，大部分黑人的经济状况比以前更糟了，尤其是年轻的黑人男性。

社会对黑人的不信任态度严重地挫伤了他们的工作积极性，尤其是年轻的黑人男性。雇主都心存偏见地认为年轻的黑人男性不够可靠，工作缺乏热情，不善于与人合作，行为也不太友善，不能很好地在一个团队里与他人合作，而且缺乏与人沟通的技巧。

不幸的是，有证据显示，即使前来求职的黑人并没有这些缺点，雇主也难以抛开这种偏见。当条件相当的黑人和白人竞争同一职位时，胜出的往往是白人。在一项研究中，研究人员让黑人大学生与白人大学生装扮成申请初级工作的求职者，结果令人极为心寒。这些年轻人都扮成穿着得体、口齿伶俐的高中毕业生，他们的资格和素质都是一样的，但即使是承认自己曾经犯过重罪[注]的白人申请者，也比记录中没有这样污点的黑人申请者得到了更多的积极回复。

如此根深蒂固的成见甚至会影响到高中毕业的黑人男性，这就意味着，在现实中学历带给黑人男性的价值要小于白人男性。不难理解，这一定会挫伤黑人男性完成教育的积极性。相较之下，黑人女性的受教育程度普遍高于黑人男性，从事的工作性质也好于黑人男性，也就讲得通了。早在1965年，也就是帕特里克·莫伊尼汉发表著名的关于黑人家庭的报告时，黑人女性高中毕业的概率就比黑人男性高出30%。2005年，在所有25～29岁的黑人中，持有大学文凭的女性和男性的比例是1.69：1。

大约在1980年，智商超过120分的黑人女性的数量是黑人男性的两倍。既然我们知道更多的学校教育和更认真地对待学习的态度能够使人变聪明，这样的事实也就很容易理解了。这种性别差异当然不是由基因造成的，因为白人男性与女性在平均智商上并不存在什么差距，而且对于白人来说，在智商较高的人中男性的数量多于女性（在智商较低的人中男性的数量也多于女性，这就是两者的平均数相差无几的原因）。如此一来，至少部分是因为学历上的优势，黑人女性在联邦政府里拥有白领工作的机会是黑人男性的两倍。

因此，社会经济地位较低的黑人不仅要面临他们这个阶层要面临的所有问题，还要面对他们的种族带来的大量其他问题。黑人的社会经济地位低于白人，而大部分黑人还背负着几乎令人崩溃的社会偏见，这些偏见不仅削弱了他们获得体面工作的竞争力，也挫伤了他们完成学业的积极性。

美国由来已久的种族歧视对黑人职业成就的影响

正如冈纳·迈尔达尔的一部经典作品的题目那样，种族问题或许是一个“美国难题”（American Dilemma），不过全世界处于社会底层的人们都面临和美国底层黑人类似的问题。根据非洲人类学家约翰·奥格布回顾的证据显示，处于社会底层的少数族裔，例如新西兰的毛利人、日本的部落民、北爱尔兰的天主教徒、以色列的瑟法底犹太人，以及在印度注定要成为社会底层的人（“不可触摸的人”），都具有类似的特点，例如学习成绩不佳，辍学的可能性极大，智力测试分数很低，违法犯罪率很高。在印度，较低种姓的人和较高种姓的人的智力差距，要比美国已知的黑人与白人的最大差距还要大。不仅如此，据记载，世界上很多非洲人以外的群体，平均智商只有85分或更低——这不仅低于过去美国黑人的平均智商，而且也远远低于黑人现在的智商。这些群体包括20世纪初在阿巴拉契亚山脉居住的白人，美国第一代意大利移民的子女，英国一些运河船工，以及苏格兰海岸附近赫布里底群岛的居民。

约翰·奥格布集中研究的是他称为“被动的”少数族裔，就像美国黑人那样被强虏后带到美国的种族。与之相反的是他称为“主动的”少数族裔，他们自

己选择离开自己的种族，就像摩门教信徒和阿米什人那样，以及所有移民至美国的少数族裔。移民往往不会将自己与其移入国家中的大多数人进行比较，他们比较的对象是自己故国的同龄人，这样一比，他们发现自己活得更好。和处于较低社会阶层的少数族裔相比，移民的犯罪率往往低于其移入国家的整体犯罪率。（不过他们的子女往往犯罪率较高，部分原因很可能是他们经过比较发现，自己的境况不如移入国家中占多数的群体。）

奥格布认为，处于较低社会阶层的少数族裔往往不能充分利用他们可以争取到的机会，因为他们不怎么相信“一分耕耘，一分收获”——也就是说，他们不相信自己的努力会得到回报。这种信念的缺失常常导致他们不努力学习，因为他们不认为努力会有所回报。少数族裔的年轻人对教育价值的看法或许与社会大相径庭。就拿美国黑人来说，年轻人会把认真学习看作是“效仿白人”的行为。奥格布几十年来一直都在研究黑人青年的价值观，尤其是黑人男青年。这些年轻的黑人学生，即使出身中等阶层家庭，念的是中等学校，也常常不做作业，并且往往更具破坏性。另外，有很多人都认为自己能够凭借体育特长去念大学，所以根本没有必要努力学习拿到高分数。即使有老师和辅导员的鼓励，他们也不会去努力学习那些较难的科目。

了解美国黑人的文化史，对了解当今的种族问题大有裨益。这里我会回顾一部分历史，并将之与另一个群体的历史做比较，即以前受到极大歧视的爱尔兰裔美国人。在19世纪的美国北方，很难预测21世纪究竟是爱尔兰人还是黑人会活得更好。在此之后，我会着重讲述托马斯·索厄尔和詹姆斯·弗林在这一方面的研究成果。

欧洲人刚到美洲大陆定居时，黑人就已经存在了。在弗吉尼亚殖民地建立12年后，20个签了契约的黑人佣工在詹姆斯顿登陆。在18世纪的弗吉尼亚，黑人与白人融合的情况与现在一样，甚至在一些方面与今天非常相像。黑人和白人去同一所教堂，如果教堂里有黑人牧师也并不稀奇。非洲文化的很多方面，包括农耕方法、饮食以及神话，都被欧洲定居者们融入了两者共同构建的文化中。

在18世纪的美国，不论是北方还是南方的黑人，虽不能说和其他群体享有平等的社会地位，但至少都得到了全社会一定的承认和尊重。很多黑人的生活和白人契约佣工的生活并没有什么两样，这些白人契约佣工在当了一段时间的学徒之后，就可以自由选择自己想做的事情，黑人也不是一辈子只属于一个主人。

从很早开始，北方的很多黑人就从未做过奴隶，或者说从奴隶制中被解放了出来。到20世纪初期，在北方的大城市里，“自由的有色人种”比爱尔兰人更容易被雇主或邻居接受。那时在工厂和商店的窗户上，还可以见到一

些标识，上面写着“黑人优先，爱尔兰人不必申请”。虽然北方大多数恢复自由身的黑人都是佣人、半技工、非技术工人，但也包括一些手工业者、小业主和早期的技术工人。在1860年的波士顿，黑人的职业地位高于爱尔兰人，纽约的酒店中黑人雇员的工资要高于爱尔兰雇员。

1855—1880年，居住在北方和南方的自由黑人家庭有85%都是由男性当家。在费城，男性当家的家庭在自由黑人中比在别的群体中更常见。而在19世纪，爱尔兰家庭中由男性当家的比例远远低于黑人家庭。直到20世纪之前，北方城市里的种族隔离情况都并不严重。1910年，芝加哥超过3/4的黑人都居住在白人占大多数的社区里。

为了理解居住在城市里的自由有色人种的生活有多么的今不如昔，我们以1800年华盛顿的500个黑人以及他们的后代为例来进行说明。这群人在1807年创办了自己的学校，在1862年被准许进入公立学校以前，黑人孩子都到黑人自己开办的学校上学。到了1870年，他们又开办了第一所黑人高中。从那时开始直到20世纪中期，该高中里3/4的学生都上了大学，甚至比今天白人上大学的比例还高。20世纪初，黑人高中的学生在全市成绩测验中的得分高于华盛顿任何一所白人高中学生的得分。最初使用智力测试时，这所黑人高中学生的得分超过了全美的平均得分。这所高中的毕业生包括美国第一位黑人将军、第一位黑人内阁成员、第一位黑人联邦法官、自南北战争后南方重建以来的第一位黑人参议员，以及血浆的发现者。

在北方的自由黑人并不是注定不能成为完全平等的公民的。他们比爱尔兰人领先一步。但是大量黑人在南方沦为奴隶，由此导致19世纪末期大量一无所有、目不识丁的黑人涌入北方城市，这才使北方黑人的努力化为了泡影。

奴隶制使得南方黑人的境遇和爱尔兰人当初在爱尔兰的境遇颇为相似。两个群体最关键的共同点就是，他们的劳动对自己而言没有创造任何价值，所以这两个群体最初并没有形成努力工作的价值观。奴隶们的劳动只是让奴隶主受益而已，爱尔兰人的劳动成果也全部归远方的英国地主所有。即使是他们居住的小屋——更多的是简陋的茅屋——都属于地主，所以爱尔兰人根本没有动力把他们的住处收拾得更好一些，因为这样只会让地主得到好处。另一个生态方面的情况也是传统的爱尔兰人不愿意工作的原因：爱尔兰的土地最适合种植土豆，但是这种作物每年只需付出几周的劳动就可以等待收获了。当19世纪中期大量爱尔兰人第一次涌入新大陆时，他们根本没有踏踏实实工作的传统。100多年过去了，他们才终于摘掉了“懒汉”的帽子。

南方废除奴隶制并没有使黑人获得经济上的自由。实际上，奴隶制下黑人

们的生活条件在很多方面都胜过南北战争后那段时期的窘况。虽然在南北战争后南方重建初期，黑人获得了少许政治自由——确实有不少黑人在当时担任联邦职务，因为曾经的南方士兵没有选举权，也不能竞选联邦职务——但是大量黑人因此不得不以租种的形式来偿还债务。到19世纪下半叶强行采取种族隔离制度之后，大部分黑人的经济情况进一步恶化，有的甚至到了令人绝望的边缘。为了摆脱贫困和歧视，南方的很多黑人都逃到了北方的中心城市地区。从19世纪末开始，出现由南方涌向北方的黑人大潮；到1940年时，每10年人数就大约会翻一番。

到达北方的黑人几乎一贫如洗，既没有掌握任何技术，也没有念过什么书，并且在生活习惯和处世态度上完全保持了乡下人的特点。他们的人数远远超过北方原有的自由黑人的后代，这改变了原来自由黑人的特征，并在北方城市中引发了大量的社会问题。

如果当时的劳动力市场完全向这些新来者敞开，那么20世纪北方城市黑人的历史或许会呈现出一番非常乐观的景象。但是总的来说，劳工运动大多将黑人排斥在外。他们不能加入工会，而那时只有成为工会会员才有可能找到高薪的工作。只有在部分地区的部分行业里，例如底特律的汽车行业，黑人才被准许大量进入工会。在其他地方，黑人们只能找到不需要熟练技术或非技术工人的工作，但即使是这样的工作，黑人也是最后受到雇用并最早遭到辞退的。

爱尔兰裔美国人因为也是白人（或者像诺埃尔·伊戈内特夫在书中写的那样，变成了白人），所以逐渐被准许加入工会，然后从社会底层逐渐升到了较低的阶层，并最终大量涌入中等阶层。爱尔兰裔美国人崛起的另外两个原因是政治——以及随之而来的资助——和天主教教会。天主教教会为了使贫穷的天主教移民得到教育，进行了英勇的斗争。

20世纪中期，在爱尔兰的爱尔兰人的智力和美国黑人的智力处在同一水平线上。英国心理学家H. J. 艾森克认为这是遗传的结果，也就是说，聪明的爱尔兰人都逃离了爱尔兰，剩下的都是呆头呆脑的爱尔兰人——还有他们的劣质基因。爱尔兰的基因库肯定比艾森克认为的更优秀，因为现在爱尔兰的人均国民生产总值已经超越了英国，儿童的读写能力也比英国的儿童强（取得这样的成就不是没有原因的，一部分是得益于从20世纪60年代开始实行的一项深入的教育举措，中学后教育的入学率从1965年的11%提高到了2003年的57%）。

虽然黑人从南方大量涌入北方打乱了北方城市原来的社会秩序，对黑人自己也产生了不利的影响（这些黑人移民几乎无法进入工会，从而找到像样的工作），但与待在南方时相比，他们的经济状况还是有了很大的改善。除了大萧条时期以外，整个20世纪中城市穷人的处境一直在稳步好转。二

十世纪六七十年代，他们的经济状况有了质的飞跃（也许并不是巧合，学习成绩取得很大进步的孩子们正是这段时期出生的）。到了1970年，夫妻双方都工作的黑人家庭与同一水平的白人家庭的收入几乎不相上下。

迈入中等阶层的黑人数量还在不断增加。拥有中等水平工作的黑人的比例，从1950年的10%上升到1976年的31%；到2002年，已经升至52%。但是早在20世纪60年代时，黑人群体的财富就开始出现两极分化。虽然中等阶层家庭的数目不断增加，但还是有大量黑人在贫困的泥潭里拼死挣扎。

黑人群体的分化与婚姻状况的稳定有着千丝万缕的联系。一个家庭内如果有男性存在，并且这个男性有工作，这个家庭的情况就会相当不错，而大约2/3的由女性当家的黑人家庭，其条件则逊色很多。有一个没有完成学业并因此经常失业的黑人男性，背后就可能有一个必须自食其力的单身母亲。每个由女性当家的家庭都有可能产生没有完成学业并因此经常失业的黑人男性，这样的男性往往也无法承担起婚姻的责任，长此以往就构成了恶性循环。

西印度群岛移民为什么能出类拔萃？

除了北方自由黑人的后代以及从前奴隶的后代，还有第三个与众不同的黑人群体。它的过去非常不同也非常多灾多难，但现在情况改善了很多，未来更将无限光明。这个群体就是美国的西印度群岛移民，他们在美国总人口中的比例不到1%，但却拥有与这一比例极不相称的卓越的非裔美国人——从黑人领袖马库斯·加维到美国第一位黑人国务卿科林·鲍威尔，都出自这个群体。1970年，第二代西印度群岛移民的平均家庭收入、受教育程度和就业比例都超过了全美的平均水平。和来自南方的黑人移民不同，移入北方大城市的西印度群岛移民和其他的外国移民群体颇为相似。他们能找到什么活就干什么活，把赚来的钱积攒起来，然后自己经营或盘下一些小本生意，他们宁愿放弃奢侈品或生活必需品，也要保证孩子受到良好的教育。西印度群岛移民从事高级职业或在职场中占据高级职位的比例，远远高于美国本土的黑人。

从表面上看，西印度群岛的奴隶制在各个方面都比美国的奴隶制更加惨无人道，因此他们能取得如今的成就更加非同寻常。西印度群岛移民在美国取得的成功并不是因为他们拥有更多的欧洲基因，或是有更多的接触欧洲文化的机会。西印度群岛移民的基因库里掺入的欧洲基因比美国黑人要少得多。和美国黑人的文化相比，西印度群岛移民的文化有着更深的非洲文化的印记（索厄尔最近认为[注]，美国旧城区“贫民窟”里的黑人文化并不是起源于18世纪的西非，而是起源于那个时期北爱尔兰和苏格兰边境地区

的文化，而后美国南方的“红脖子”乡下人把这种文化的火炬传给了南方的黑人）。

在索厄尔看来，西印度群岛移民之所以出类拔萃，关键在于新大陆上的奴隶制经济史。在美国南方，奴隶们都是在一间厨房里一起吃饭，或者领到食物后回到自己的住处再进一步加工。而在西印度群岛，黑人们自己种庄稼，然后把剩余的粮食拿到市场上出售。由于白人很少，因此所有的技术劳动、手工艺活以及一大部分小买卖都不得不由黑人独立承担。虽然西印度群岛移民初到美国时很穷，但是和他们的南方黑人兄弟相比，他们能够更好地把握美国的种种机会。

但我必须指出的是，西印度群岛移民并不是从全部西印度群岛人群里随机挑选的代表。西印度群岛移民中没有包括多少在西印度群岛社会最底层的人，这些移民中专业人士和管理阶层所占的比例，大大高于西印度群岛人口的整体水平。简而言之，造就西印度群岛移民在美国高成功率的一部分原因就是自我选择，而这种自我选择无疑和他们的优秀智力基因有一定关系。

虽然西印度群岛移民中，具有一技之长的人的比例要比平均水平高，但他们的学业和职业成就告诉我们一件非常重要的事情，那就是种族主义是造成黑人职业成就较低的罪魁祸首。如果黑人掌握过硬的技术，并且对工作抱有积极的态度，那么不管种族主义多么严重，也不能阻止黑人取得很大的成就（当然，在纽约和其他西印度群岛移民大量聚集的地方，人们对这个群体的固有印象是他们很勤奋，并且值得信赖，这一点帮了他们不少的忙。抑扬顿挫的加勒比口音在劳动力市场上可是很吃香的）。

养育方式的差异会影响孩子认知能力的发展

西印度群岛移民在文化上投入的资本高得不同寻常，而有些本土黑人在文化方面的投入明显低于白人。在所有美国黑人整体上具备的劣势之中，首当其冲的就是许多黑人父母养育子女的方式。与社会及经济状况都较为接近的白人相比，用他们的养育方式教育出来的孩子，不太可能取得智力测试的高分以及优异的学习成绩。

没有什么比许多黑人父母和子女进行交流的方式更能说明问题的了。我在上一章中曾指出，精英阶层的孩子每天能够听到2 000个单词；而工人阶层的孩子，不管是黑人还是白人，每天仅能听到1 300个单词。靠救济维生的黑人的孩子每天听到的单词量只有600个。精英阶层的孩子到三岁时，已经听到过3 000万个单词，而父母靠救济维生的黑人孩子只听过1 000万个单词。与靠救济维生的黑人妇女与子女交谈时所使用的词汇相比，精英阶层三岁大的子女与父母交谈时使用的词汇要丰富得多。

让我们回想在上一章中提到的人类学家雪莉·布莱斯·希斯所做的关于美国南方一个农村社区里的孩子的研究。整个20世纪60年代和70年代，希斯和三种不同类型的家庭共同生活达数月之久，并跟踪记录了孩子们在学校的表现。她研究的家庭包括父母至少有一方是教师的中等阶层白人家庭，工人阶层的白人家庭以及社会阶层较低的黑人家庭。工人阶层的家庭在很多方面不同于中等阶层的家庭，而这种差异在不少穷苦的黑人家庭里也比比皆是。不过，在语言的社会化方面，黑人既不同于工人阶层的白人家庭，也不同于中等阶层的白人家庭——他们的养育方式甚至远远不足以使子女做好上学的准备。

穷苦的黑人孩子往往出生在一个人口众多的大家庭里，家里人与孩子既有口头上的沟通，也有非口头上的。大人们通常都不直接跟孩子讲话——他们懒得去猜孩子发出的声音究竟代表着什么意思。他们不会使用适合儿童的语言和孩子说话，他们也不会赋予物体或事件特定称谓，或者试着将当下的物体与在其他环境中遇到的其他物体联系起来。换句话说，他们不会将事物串联起来，使孩子学会触类旁通。

这些黑人孩子也没有什么能够开发智力的玩具，他们的玩具只是些不会伤着人的家居物件——勺子、塑料餐具，还有壶盖什么的。大一点儿的孩子或许有一些电动或机械玩具。不过他们都没有锻炼动手能力的玩具（积木、能够拆装的玩具或者拼图）可以玩。

他们也没有什么书可以读。大人们读报纸、邮件、日历、广告册子以及《圣经》，但是却没有为孩子准备专门的儿童读物，除了有时从主日学校拿回的一点材料外。大人们也不会耐心读故事给孩子听。晚上临睡前也没有什么惯例活动，甚至入睡前父母也不给孩子们讲睡前故事。

而且，这些父母也不会就孩子周围的环境提出“什么”式的问题，他们问的都是不够具体的比较，例如“那个像什么”。（这很有可能会提高孩子们辨别相似之处的能力。黑人在智力测试的类同小测验中的成绩是所有小测验里最高的。）这种将两个事件以隐喻的方式联系在一起的能力，并不会给孩子们在学校里的生活带来优势。实际上，这种能力经常会造成问题，因为孩子们看到的联系并不是老师所要求他们指出的那种联系。到了高一点儿的年级，孩子们要学习比较和评估的时候——也就是孩子们发现相似之处的能力本可以大展拳脚的时候——这些孩子就远远地落在了后面，因为他们不具备口头及书面的理解能力，无法找出老师可能认可的那种相似点。

在家里，父母们的问题也不是已知答案类型的问题——也就是说，那些大人知道答案的问题（“大象是什么颜色的，比利”）。这样一来，孩子们上学后根本不知道应该如何回答这类问题。即使老师提出最简单的问题，他

们都有可能答不上来，因为这种提问的方式使他们困惑（“要是连老师都不知道答案是什么，我肯定也不知道”）。

不过，孩子们在家里有时还是会讲些故事的——如果他们能够长时间地吸引家庭成员注意力的话。但是以学校的标准来看，他们讲的故事可并不怎么样：这些故事通常都没头没尾——只不过是为了娱乐一下，听众听了一会儿就失去兴趣了。这些孩子的叙述能力超过了工人阶层的白人孩子，甚至胜过很多中等阶层的白人孩子。（黑人的叙述能力从娱乐业和神职人员身上就可以看出来。有一句话说得好：最差的黑人传教士礼拜日在教堂里的最差表现，也要好过最好的白人传教士礼拜日在教堂里的最好表现。）

在学校里，老师要求孩子们根据形状、颜色或大小将物体分类。但是贫穷的黑人孩子根本没学过怎么分类，他们对这样的练习非常陌生。不仅如此，如果老师要求他们口头描述书中出现的场景时——这些场景都有固定的描述模式——他们通常很难将这些场景与现实世界里具体的物体或事件联系起来。

20世纪80年代末，希斯又回到了北卡罗来纳州，继续研究20多年前她所研究的那群孩子的子女。以前她住过的社区已经不复存在了，原来一些家长工作的那家纺织厂也早已关门了，农场实现了机械化。当年的孩子们中有一部分现在已经在城市里找到中产阶层的工作，剩下的人则是希斯这次研究的对象，他们有的依旧居住在原来那个城镇的贫民窟里，有的则住在亚特兰大一幢低收入者聚集的高楼里。

当年的那些孩子现在都已经为人父母，并且通常都很早就结婚生子了。实际上，希斯当年研究过的社区里，每一个女孩都在十几岁时就已经有了一个孩子。这些年轻的妈妈既不会使用适合孩子的语言与孩子交流，也没有告诉孩子环境中出现的东西的特定称谓。她们也从不让孩子们描述每天是如何度过的，而当孩子开口说话时，母亲有时候还会嘲弄孩子们卖弄知识。

在大多数情况下，孩子都交由她们的妈妈来抚养，这些十几岁的年轻女孩则返回高中念书，恢复自己的社交生活。希斯的研究对象中有一个十几岁的女孩，大多数时间都靠救济度日，她很少与自己的孩子进行交流，或是对孩子说上几句话，即使开口对孩子说话，一般也不会超过一分钟。与她所在的那个北卡罗来纳州的大家庭相比，她的语言世界和社会生活要单调许多。她很少主动开口说话，只是被动地收看电视节目，或是读一些电影或电视方面的杂志。与她来往的人大都是她所在的住宅区内那些做临时工的妇女。这些女性只盯着眼前的事——怎样准备好材料去福利部门申请救济，或是托儿所的规定可能会发生变化，又或是怎样才能够让孩子的父亲重新开始寄孩子的抚养费过来。

从希斯的报告中不难看出，20世纪80年代末出生的年轻父母生育的孩子中，至少有一部分孩子的认知、社会及情感生活远不及他们的父母成长时那么丰富多彩。

另外一项关于20世纪80年代典型黑人家庭生活的研究更加系统，同时也印证了希斯的观点，那就是，在那个时代以及现在，黑人的家庭环境不仅不利于智力的发展，而且情感上也不够温暖。梅雷迪思·菲利普斯、珍妮·布鲁克斯—冈恩以及他们的同事对一些研究的结果进行了详细的分析，这些研究评估了黑人家庭与白人家庭中各个方面的家庭环境情况。他们的分析建立在两组数据的基础上。一组数据来自一项叫作“美国青少年研究调查（儿童组）的研究。这项研究始于1986年，研究对象是在1979年年龄介于14～22岁的青少年的子女，共计6 000个孩子。这项研究考察了大量人口变量和家庭变量。另一组数据来自婴儿健康和发展项目。这个项目研究了在8家不同医院出生的婴儿，这些婴儿出生时的体重都不足2.5千克——这样的婴儿不仅可能智商低，而且可能还存在不少身体健康方面的问题。有一项干预计划旨在提高这些孩子的智力和学习成绩。在下一章中，我会谈谈这个耗费了大量的时间、物力和财力的计划所取得的成效。不过现在我们还是只讨论这个项目控制组中的315名黑人和白人孩子，这些孩子出生时的体重为2～2.5千克。

菲利普斯及其同事考量的主要指标，基本上就是家庭环境测量观察法（HOME，Home Observation for Measurement of the Environment）所采取的测量标准。采访者到研究对象的家中进行观察，并向孩子的母亲询问一些情况，然后给出这个家庭在考量指标上的得分。这些指标包括“从家庭外部获得的学习经验（参观博物馆、拜访朋友、去杂货店买东西），从家庭内部获得的文化经验（孩子拥有10本以上的书籍、母亲读书给孩子听、家庭成员读报纸、家庭订阅杂志），家庭内部的感知刺激活动（能够促进技能学习的材料，例如辨别字母、数字、颜色、形状以及大小的能力），惩罚（在采访者家访的过程中孩子是否挨过打以及母亲管教孩子的风格），母爱（在家访过程中母亲是否亲吻、抚摸或拥抱了自己的孩子，在家访过程中母亲是否称赞了孩子的成就），以及其他物理环境因素（家里是否足够干净整洁，孩子游戏的环境是否安全）”。

黑人家庭和白人家庭在这两项研究中显示出极大的差异，在某些考量指标上，两者的差距竟然高达0.6个标准差。黑人家庭在考量指标上的得分往往都与认知变量密切相关。在美国青少年研究调查（儿童组）的研究中，和从来不读书给孩子听的情况相比，如果每天读书给孩子听，五六岁孩子在词汇上的得分相当于孩子的智商提高4分。在婴儿健康和发展项目中，如果一个家庭的得分比所有考量指标得分的平均值高出1个标准差，那么这个家庭里的孩子的智商就比别的孩子高出9分。

我在第5章中曾经提到过哈特和里斯利有关堪萨斯城的家庭的研究。哈特和里斯利发现，来自不同群体的父母在对待孩子的方式上（是慈爱还是严厉）存在着巨大的差异。让我们回想一下，身为精英阶层的父母每责备孩子1次，就有6次鼓励；一些工人阶层的父母每责备孩子1次，只会鼓励孩子2次；而靠救济生活的黑人父母，每责备孩子2次，才会鼓励1次。到3岁时，精英阶层的孩子受到了50万次的鼓励和8万次的责备，而母亲靠救济生活的黑人孩子则受到了7.5万次的鼓励和20万次的责备。

我们完全有理由相信，这样的差异会对孩子认知能力的发展产生重要的影响。我们也知道，在一定程度上，黑人父母养育方式中某些不利的方面在中等阶层的黑人家庭中也很典型，至少在20世纪80年代是这样的。让我们回想一下埃尔茜·穆尔关于中等阶层白人或黑人父母抚养黑人或混血孩子的研究。白人养父母收养的黑人或混血孩子的智商，比黑人养父母收养的黑人或混血孩子的智商高出13分。虽然我们不知道这样的差距究竟有多少是家庭环境造成的，多少是学校和邻里环境造成的，但是有一点我们是肯定的，那就是黑人孩子成长的环境不会像白人孩子成长的环境那样能促进孩子认知能力的发展。

穆尔和研究人员走进每个受访家庭，让孩子完成一项积木图案的任务——也就是让孩子先看一个图案，然后尝试用积木把这个图案摆出来，孩子的母亲在一边观看。和黑人养母相比，白人养母对孩子的鼓励远远多于责备。在摆积木的过程中，当孩子们遇到问题时，白人养母往往会开开玩笑，以缓解孩子的紧张；而黑人养母更多地会皱眉或沉下脸来。白人养母会鼓励孩子努力解决问题（“哎呀，这真是个有趣的想法”或者“你可真在行”），而黑人养母则更多地表示否定（“你知道那样摆肯定不对”或者“如果你真努力了，肯定比现在做得好”）。白人养母还会向孩子们提供帮助，而她们的方式会促进孩子们学习，例如她们会提出建议，孩子们会利用这些建议自己探索出如何将积木摆成要求的图案（“你为什么不试试每次只摆一部分呢”）；黑人养母则会给出具体的指示，这样孩子们就没有机会自己去发现如何完成这项任务（“应该用那块，不过你得先把它像这样翻转过来”，然后示范正确的方法给孩子看）。白人养母的态度基本上都是“只要你努力了，出错也没关系”；而黑人养母则往往会发火，并且对孩子缺乏信心。

对于这些有关白人养母和黑人养母行为的证据，我想强调两点：第一，我们不清楚阻碍孩子智力发展的究竟是上述的行为，还是诸如周边环境、同龄人的影响或者学校等其他环境因素。第二，如今不太可能在中等阶层的白人养母和黑人养母之间再发现这样的差异了。这项研究差不多已经过去25年，第二代中等阶层父母的养育方式可能不同于第一代中等阶层父母。当然我们可以期待，第二代中等阶层父母的行为方式会促进孩子的探索能

力和智力发展。

上一章的结论与这一章的结论是契合的。黑人与白人的智力差距不是基因造成的，可测量的环境因素才是造成差异的全部原因。虽然较低社会阶层的黑人和白人同样地受到许多不利条件的影响，但是黑人——尤其是工人阶层和处于社会底层的黑人——还要忍受种族歧视，而正是种族歧视阻碍了他们的职业发展。与白人文化相比，黑人文化中的某些方面——存在于每个社会阶层的黑人家庭中——会阻碍孩子认知能力的发展。除了中等阶层的黑人以外，其他阶层黑人面临的周边环境和选择的学校，不仅使这些文化的不利影响扩大了，还使这些黑人很难摆脱这种文化的影响。即使是中等阶层的黑人，在青少年时期也受着否定成就的亚文化的影响。在这种亚文化的鼓动下，这些青少年往往认为体育特长、表演天赋和在街头闯荡的本事足以代替学习能力。

我相信，按照正常的发展进程，黑人的社会经济状况和智力水平会慢慢得到提高。在过去的几十年里，黑人的犯罪率和吸毒率持续下降（不过自从2005年开始，暴力犯罪率又开始有所回升）。不仅如此，还会有越来越多的黑人跻身中等阶层，黑人的智力和学习成绩也会逐步提高。

在下一章里，你会看到怎样做才能使穷苦的黑人迈入工人阶层，工人阶层的黑人又如何逐步迈向中等阶层。

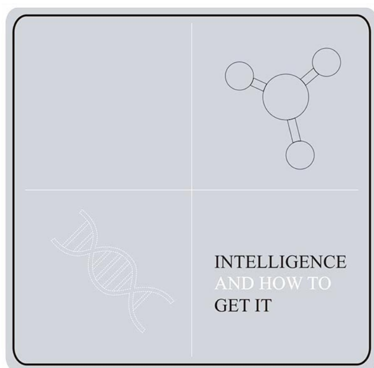
---

1. 摩尔人（Moors），中世纪时西欧西班牙人和葡萄牙人对北非穆斯林的贬称。摩尔人实际上是阿拉伯人、柏柏尔人以及北非黑人的混种人，公元8世纪皈依伊斯兰教并由北非进入并统治西班牙等地。——译者注

2. 美国各州对重罪的定义不尽相同，但大致上都是以暴力为手段的犯罪，包括谋杀、强奸、恶意伤害和侵入私宅等。——译者注

3. 这一点值得我们详细引用索厄尔的观点。下面这段话来自他的著作《狭隘的黑人与开明的白人》的第6页，“在南方白人之间大行其道的文化观和社会模式包括：厌恶工作，倾向使用暴力，不重视教育，私生活糜烂，缺乏远见，酗酒，不思进取，不顾后果地寻求刺激，喜欢奔放的歌舞，他们的宗教演讲往往充满过于华丽的词藻，激情洋溢，联想丰富”。下面这段话来自第1页和第2页，“这种文化很久以前就在它起源的英国消亡了，不过在美国南方却存活了下来。后来，在南方黑人和白人之间也几乎销声匿迹了。但是今天，在最贫穷最落后的城市黑人贫民窟里，还是可以看到这种文化的影子”。

第7章
消除智力差距是痴人说梦吗？



我们已经尝试过补偿教育了，不过根本没有用。

——阿瑟·詹森（1969年）

没有证据显示学校改革能够显著缩小“能力”测试所反映的认知能力的差距。

——克里斯托弗·詹克斯等

没有理由相信“显著并永久性地提高智力”应该作为我们当前的一项政策，这和我们愿意花多少钱没有关系。

——查尔斯·莫瑞（2007年）

2002年，美国国会通过了《不让一个孩子落后法案》，要求美国学校在2014年以前消除不同社会阶层以及少数族裔和白人在学习成绩上的差距。我不清楚大多数国会议员是否真的相信如此辉煌的成就是可以实现的。如果他们确实相信的话，那么只能说他们对那些能够成就优秀学业成绩的因素简直一无所知。

刺激孩子的智力发展，支持孩子在家里、社区和学校进行探索并取得成就，这才是智力资本得以积累的方式。认为通过命令——而且仅在学校里

实行——就可以提高智力的看法，简直就是天方夜谭。不仅如此，少数族裔和穷人就读的学校在很多方面都迫切需要加以改善，而这些问题都不是一朝一夕能够解决的。这些问题包括，愿意在这些薪酬不高的学校里任教的老师的素质、学校管理的质量、高年级学生不服从老师管理扰乱教学秩序、学生的家庭以及他们居住的社区都不太可能鼓励他们追求优异的学习成绩等等。

看过前一章后我们就应该清楚，黑人与白人在学习成绩上的差异最终能够得以消除的论断，在理论上完全成立。我们通常所说的“拉美裔美国人”也是一个人数众多、高度多元化的文化群体。虽然能够证明他们天生的智商并不低于白人的证据少之又少，但我相信他们与白人的差距有朝一日一定能够消除。

另一方面，我们还要明白，与黑人/白人和拉美裔美国人/白人在成就和智商上的差距不同，社会阶层之间的差距是永远不能够消除的。这毋庸置疑，没有什么特别的原因，只是因为富人总是想方设法让他们的子女接受更好的教育，因此总能在养育方式上更胜一筹，并且总能为子女提供一流的居住环境。此外，较低社会阶层和中等阶层的基因构成多少还是有些差别的。回想一下第1章，在一个家庭里，智商较高的子女的社会经济地位要比其他智商平平的兄弟姐妹高得多。既然高智商部分得益于更优秀的智力基因，那么较高的社会经济地位也总有一部分是得益于优质的智力基因。所以社会经济地位较高的人能够把自己更优秀的智力基因遗传给后代，并且为后代提供更加有利的环境。

不过我们不应该因此就抱有悲观的看法，认为社会阶层较低的人的智力无法得到较大的提高。回想第2章（关于遗传度的），社会经济地位较低的人的子女如果在中上阶层的家庭长大，其智商可以提高12～18分。社会经济地位较低的人想要提升智力，确实大有空间。

对于社会经济地位较低的人和目前处于不利地位的少数族裔，我们能够期待他们在多大程度上实现智力的提升呢？

儿童早期教育与智力改善

当我告诉别人我正在写一本关于智力可变性的书时，有些人会告诉我，美国先行教育计划没有效果，劝我别白白浪费时间。对于很多人来说，他们对于某一项教育计划的看法就预示着他们对智力能否发生改变的看法。

美国先行教育计划是一项补偿计划，最初主要是为了改善三到四岁贫困儿童的健康和福利状况。这项计划的开创者希望该计划还可以提高孩子们的智力和学习成绩，并使这些孩子在未来的生活中取得更大的成就。该计划

每次活动半天，每周5次，通常持续34周，不过每次活动只有一小部分的时间专门用来训练孩子的认知能力。

美国先行教育计划是个失败的计划吗？这就取决于你如何看待这个问题了。从改善身体健康状况的角度来看，这项计划取得了十足的成功。相比未参与该计划的健康状况相近的儿童，参与这项计划的儿童的死亡率要低33~75个百分点。实际上，参与这项计划的儿童的死亡率已经降到了美国儿童的整体死亡率水平。

在这项计划实施的早期，完成该计划的5岁儿童在认知测试中的得分提高了大约0.35个标准差，相当于智商提高了5分。而最近的研究显示，孩子六七岁时某些智商和成就变量上仍然保留着0.1~0.2个标准差的水平。不过到了孩子升入小学高年级时，这些效果就完全消失了。最近有一些报告显示，5岁的效果值更低——接近0.25个标准差。不过在那个时候，对学前教育措施进行效果测试时，很难找到真正意义上的控制组，因为那时即使是穷困的少数族裔的孩子，也有一大半在上学前班之前就已经接受过一定的教育了。这样一来，在比较效果大小时，比较的就不是治疗组和完全没有接受过治疗的控制组了，而是与其中半数孩子都上过日托班的控制组在做比较。

令人吃惊的是，几乎没有什么评估美国先行教育计划对孩子学习成绩产生的长期效果的研究。仅有的一些研究显示，孩子们念完高中时这项计划还保留极少的一部分效果，比控制组大约高出2~5个百分点，当孩子们即将升入大学时还有一小部分成效在，比控制组大约高出3~6个百分点。美国先行教育计划的成本大约是每个孩子7 000美元，所以，从提高智力和学习成绩的角度来看，这样的回报究竟值不值得，恐怕每个人都有不同的答案。

早期先行教育计划从婴儿刚出生时就开始实施，到孩子3岁时结束。不过和先行教育计划一样，这项计划也被证明不会改善孩子的智力状况。这项计划的服务项目包括儿童成长、儿童看护、家访、父母养育方法教育，以及家庭支持服务。每个家庭都可以根据孩子的实际情况自由掌握各项服务所占的比例。在考察这项计划的效果时我们衡量了一系列变量，从纯粹与认知有关的变量到与情感和社会有关的变量，这些变量的效果值介于0.1~0.3个标准差之间——对少数族裔孩子产生的效果略大于对白人孩子产生的效果。取得的最显著的短期效果也就是智商提高了不到4分而已（尽管词汇得分提高了0.4个标准差）。这项计划花费不菲，但是现在仍无法确定其长期效果（很可能不会得到测量）以证明这样的投入是值得的。

还有比先行教育计划更耗时耗资的计划，并且其中一部分计划取得的效果

更大更持久。一些托儿所和学前班项目主要面向黑人儿童，这些项目的规模要小一些，但执行的情况也相对好一些。对十几个这样的项目进行研究后发现，这些项目大大提高了孩子们的智商——在孩子5岁时提高了0.7个标准差甚至更多。不仅如此，孩子们在小学低年级时的学习成绩也取得了很大的进步。不过这样的效果一般都会逐渐减弱，最后完全消失。

如果无法维持高质量的环境，这种成效尽失的结果是在所难免的。只有孩子们的大脑像黏土一样，我们才能期待大脑一旦被塑造成形，在很多年以后仍能保持当初的形状。而如果孩子们的大脑像肌肉一样，那么只有勤加练习，例如刺激智力发展的环境和活动，才能使其保持良好的表现。我倾向于认为孩子们的大脑更有可能像肌肉一样，数据显示的结果也是如此。

实际上，一些儿童早期教育计划确实使智商在短期内得到了大幅提高，在长期内也使智商或学习成绩得到了显著提高，有些孩子的智商和学习成绩甚至都得到了提高。让我们来看一看更加有效的三项计划。这些计划随机选取了一些孩子编入治疗组和控制组，然后一直跟踪记录到这些孩子即将成年或者成年以后的数年时间。

1962—1967年，劳伦斯·施瓦哈特和戴维·韦卡特在密歇根州的伊斯兰提开展了一项叫作佩里计划的早期教育计划。总共有58名黑人儿童参加了这项计划，这些孩子家境贫困，母亲在斯坦福—比奈智力测试中测得的智商在75～85分之间。计划开始实施的第一年里，孩子们参加时的年龄是4岁；后面的4年里，孩子们参加时的年龄是3岁。除了第一组孩子，其他孩子参加这个计划的时间都是两年。

佩里计划要求在一年的时间里，孩子们每天上午在教室里度过两个半小时，持续30周的时间，主要是进行一些促进孩子认知和社交能力发展的活动。孩子与老师的比例平均是6：1，这个比例确实很低，所有的老师都接受过很多儿童早期发展和教育方面的培训。每周老师会进行一次家访，在孩子家里待上90分钟，鼓励孩子的母亲参与到孩子的教育过程中来。参与这项计划的58名孩子构成了治疗组，其他65名孩子构成了控制组。当项目计划结束后，这些孩子都进入他们所在社区里教育水平较低的社区学校上学。

项目结束时孩子们5岁，控制组中孩子的平均智商是83分，治疗组中孩子的平均智商是95分。此后，治疗组孩子的智商随着年级的增加逐渐下降，到10岁时，他们的智商已经下降到了控制组的平均水平，也就是85分。

这样的结果令人失望，但考虑到孩子们在项目结束后的家庭情况、居住环境和就读的学校，这一点也就不足为奇了。不过令人感到吃惊的是，孩子们的学习成绩和后来的社会经济地位都取得了很大的进步。这些都显示在

了图7–1（a）中。控制组的孩子大约有1/3在某个时期内不得不接受一些课外辅导，而治疗组只有13%的孩子接受过课外辅导。14岁时，控制组里只有14%的孩子在加州成就测验中的百分位排名高于第10位，而治疗组里有大约一半的学生的百分位排名都高于第10位。阅读、数学和语言成绩的得分比控制组高出0.5~0.75个标准差。控制组里有43%的孩子能够从高中毕业，而治疗组的这一比例是65%。治疗组的高中成绩比控制组高出0.57个标准差。到27岁时，控制组里只有6%的人月工资达到2 000美元，治疗组里则有28%的人可以挣到这样的工资。控制组里有11%的人拥有自己的住房，而治疗组里有33%的人拥有自己的住房。控制组里有20%的人从未接受过救济，而在治疗组里这一比例达到了40%。控制组里有8%的女性结婚，而治疗组里有40%的女性结婚。到了40岁，控制组里有55%的人被拘捕了5次以上，而这一比例在治疗组里是36%。这样的成果对社会和经济都产生了重大的影响。

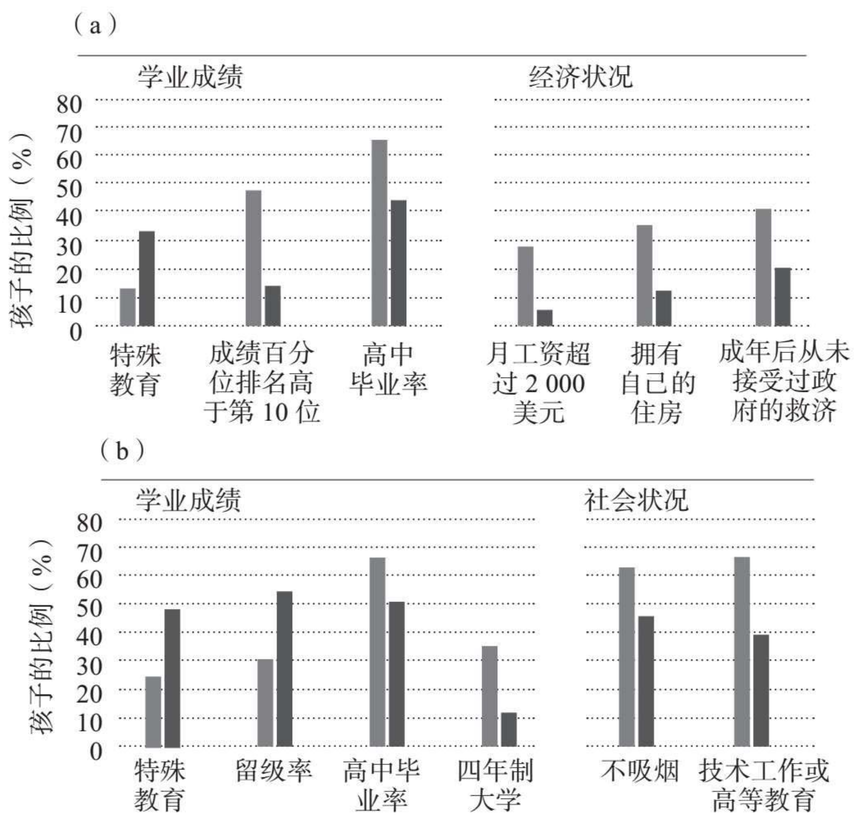


图7–1 佩里学前教育计划和初学者计划的参与者在学习、经济和社会方面的成果*

* 浅色的图柱代表治疗组，深色的图柱代表控制组。
（a）佩里计划在参与者年满27 岁时收集到的数据。成绩百分位排名高于第10 位，是指孩子们14 岁时在加州成就测验（1970 年）中的排名高于成绩最低的10% 的学生；高中毕业率，是指按时从高中毕业的孩子的比例。
（b）初学者计划在参与者年满21 岁时收集到的数据（卡罗来纳初学者计划以 及卡罗来纳反应式教学方式，1972—1992 年）。
资料来源：Knudsen, Heckman, Cameron, and Shonkoff, 2006.

不光是佩里计划，还有一些教育计划也显示，虽然治疗组的孩子们获得的智商优势逐渐消失了，但是后来他们在生活中取得了更多的成就。这些成就包括留级率和接受课外辅导的比例降低，高中毕业率升高。即使智商上的优势完全消失以后，学习成绩和生活状况还能得到如此大的提高，这使得很多人怀疑这样的提高恐怕不是由于智商提高本身造成的，而是因为这样的干预措施导致孩子发生了一些性情上的变化，或积极性得到了提高。即使当环境不再有利于提高智力时，这些变化也能使得孩子们取得更大的成就。

一群研究人员在密尔沃基实施了一项比佩里计划更加耗时耗资的早期教育计划。计划的发起人里克·希伯发现，这个城市里有一个很特殊的地区，这一地区的人口占全市总人口的3%，但智力发展迟缓的儿童的数量却占到全市总人口的33%。于是他决定把这个项目的资源集中用到这一地区。参与这项计划的孩子都是非裔美国儿童，这些孩子都是智力发育迟缓的高风险人群，因为他们的母亲经济条件很差，并且智商只有75分甚至更低。研究人员随机选取了一些孩子编为控制组（18个孩子）和治疗组（17个孩子）。编入治疗组的孩子要参加该计划所提供的日托项目，从不满6个月时开始，一直到上小学一年级之前结束。

密尔沃基计划的目的是给孩子们提供类似中等阶层家庭的环境，重点培养孩子们的语言能力和认知能力。辅导人员经常利用当时最好的发展计划和益智玩具与孩子们进行有趣的互动。每天的活动时间是7小时，每周5天。这项计划不仅为孩子们提供了营养充足的食物和高品质的医疗及牙齿健康服务，而且还对孩子们的母亲进行家政和幼儿保育方面的培训。研究人员不仅常常将治疗组的孩子与控制组的孩子进行比较，还把他们与另一个比较组里的孩子进行比较。这个比较组里的孩子几乎不存在智力发育迟缓的风险，因为他们的母亲的智商都达到了中等或中上水平（使用韦克斯勒成人智力测验量表测得的智商是108分）。在孩子们30个月大的时候，使用斯坦福—比奈智力测试测量他们的智商，控制组孩子的平均智商是94分，治疗组孩子们的平均智商是124分，比较组孩子的平均智商是113分，也就是说治疗组孩子们的平均智商实际上超过了比较组的孩子。到孩子们5

岁时，使用韦克斯勒学前儿童及小学生智力测验量表测量他们的智商，控制组孩子的平均智商是83分，治疗组孩子的平均智商是110分，而比较组孩子的平均智商是101分，治疗组孩子的平均智商仍然高于比较组里那些低风险的孩子。

当孩子们7岁时，也就是这项计划结束的时候，治疗组的平均智商仍然高出控制组22分。（请注意，这一差距甚至大于在较低社会阶层和较高社会阶层的家庭中长大的孩子之间的智商差距。）到了要进入小学学习时，与佩里计划中的孩子们不同的是，密尔沃基计划的孩子进入了相当优秀的学校——这些学校在成就测验中的水平都达到或超过了全美平均水平，这使得治疗组里的孩子保持住了已取得的智力优势。在这项计划结束9年后，也就是孩子们到青少年时，使用韦克斯勒儿童智力测验量表对他们的智商进行测试，控制组的平均智商是91分，而治疗组的平均智商是101分——与比较组大致处于同等水平。比较组里那些母亲拥有中等或中上水平智商的孩子，平均智商是97分。

从小学一年级到四年级，治疗组的学习成绩一直高于控制组，如果以标准差来计算的话，两者的差距很大——大约0.75个标准差。不过，由于样本的人数较少，所以就存在10%的可能性，即如此大的差距可能是偶然获得的（虽然对于任何一组来说这种差异的可能性都是10%，但是单就治疗而言，被高估的可能性只有5%）。

克雷格·雷米和弗朗西斯·坎贝尔及其同事尝试了一项更为细致复杂的计划——初学者计划。有111个孩子参加了这项计划，这些孩子出生于1972—1977年之间，绝大多数都是非裔美国儿童。这些孩子被认为是面临极高的智力发育迟缓的风险，因为孩子的母亲平均智商只有85分，受教育年限平均只有10年，还有其他一些风险因素，例如家庭收入低、母亲独自抚养孩子、母亲缺少社会救助和家庭支持、兄弟姐妹学习成绩差、父母从事非技术类工作、依赖公共机构的救助等。这项计划是一项全日制干预计划，全年没有一天中断，从孩子们未满6个月时开始实施。刚开始时幼儿与教师的比例是3：1，后来随着项目的进行提高到6：1。这一干预活动一直持续到孩子们进入学前班学习为止。此后研究人员坚持定期收集参与者的数据，直到他们年满21岁。

研究人员把孩子们编为4组，每组大约25个孩子。其中一组不但参加学前干预计划，还参加学龄儿童干预计划。在孩子念小学一年级到三年级期间，学龄儿童干预计划会派出一名家庭—学校教师，这位教师经常和孩子的父母见面，教给他们如何在家里进行辅助的教育活动，并鼓励家长每天至少抽出15分钟的时间和孩子一起学习。家庭—学校教师还是家长和孩子的老师之间沟通的桥梁，她们每两周和家长以及孩子的老师碰一次面。此外，她们还帮助这些家长找工作，和社会服务机构打交道，并带孩子去看

医生。除了第一组孩子外，另外一组孩子只参加了学前干预计划，一组孩子只参加了学龄儿童干预计划，第四组孩子作为控制组没有参加任何干预计划。

到孩子们3岁时，没参与学前干预计划的孩子平均智商是84分，参与了学前干预计划的孩子平均智商是101分。到学前干预计划结束时，控制组的平均智商是94分，参与了学前干预计划的孩子平均智商是101分。接着，孩子们并没有进入位于市中心平民区的水平较低的小学，而是进入了大部分都是家境富裕的白人孩子的学校读书。到孩子们12岁时，参加干预计划的孩子中只有13%智商低于85分，而在控制组中这一比例高达44%。即使到了21岁，参加学前干预计划的孩子的平均智商，也比没有参加学前干预计划的控制组里的孩子高出4.5分。那些母亲智商最低（不足70分）的孩子从这项计划里受益最大。不过证据显示，学龄儿童干预计划并没有进一步提高接受学前干预计划的孩子的智力水平，学龄儿童干预计划本身也没有取得多大成果。这项研究导致很多人对家访的效果心灰意冷——除非是投入了很多心血和时间的家访。

等研究对象到了21岁，就可以明显看到，初学者计划中的学前干预计划对学习成绩产生了极大的影响，这些影响都列在图7–1（b）中。控制组中几乎一半的孩子在学习生涯里的某段时间都接受过课外辅导，干预组里只有不到1/4的孩子接受过课外辅导。控制组中超过一半的孩子曾经留过一级，而干预组里只有30%的孩子曾经留过级。在15岁时，干预组孩子的阅读得分比控制组孩子的得分高1.4个标准差，数学得分高0.86个标准差。到21岁时，干预组孩子的阅读成绩比控制组的孩子整整高出两个年级的水平，数学成绩高出不止一个年级的水平。控制组的孩子有半数从高中毕业，干预组里有2/3的孩子从高中顺利毕业。控制组的孩子只有12%继续读了四年大学，而干预组里有1/3的人都念了4年大学。在21岁时，干预组里有2/3的人得到了一份技术工人的工作或在大学里深造，而在控制组里这一比例不足40%。

不过初学者计划的实际效果很有可能比这样的成果还要大一些，因为控制组里的孩子都曾经接受过一些学前教育。

有几项研究对初学者计划的一部分进行了重复测试。其中一项叫作关爱计划的研究比较重要，因为这项研究的结果可以用来驳斥赫恩斯坦和莫瑞的观点。赫恩斯坦和莫瑞认为，初学者计划得到的结果并不可信，因为他们发现，参与者一岁时的智商就存在很大的差距。那时这项计划并没有实施多久，所以计划本身不可能造成这样大的差距。这就表明，这不是计划取得的成果，而是研究人员从一开始就未能完全随机选定智商相当的控制组和干预组。关爱计划使用了和初学者计划基本相同的方法，在编定控制组和干预组时，选择了6个月大时在贝莱智力发育指数中得分完全相同的孩

子。等到孩子们一岁时，智商得分相差11分。所以初学者计划对如此小的孩子的智商也可以起到很大的作用。

另外一项重复测验尤为重要，因为根据这项研究的结果，那些因早产或出生时体重过轻——2.0～2.5千克——而面临智力发育迟缓风险的孩子，也可以通过初学者计划来提高智商。研究中的婴儿有2/3都是来自黑人家庭或拉美裔美国人家庭。这些孩子的母亲，有大约40%高中没有毕业，只有13%读完了大学。这项计划和初学者计划的不同之处在于，这些计划不是从孩子出生不久开始的，而是从1岁时开始3岁时结束的。孩子们在3岁时分别参加了几种不同的智力测试，结果显示，干预组的平均智商比控制组高出9.2～12.5分。

这项计划结束两年后，根据使用的智力测试，干预组的平均智商比控制组高出2.5～5.4分（2.5分在统计学意义上并不显著）。孩子们8岁时，这项计划仍然还有一些效果，根据使用的智力测试，干预组的平均智商比控制组高出3.6～5.4分。即使是在18岁时，仍然可以察觉到这项计划的效果，干预组的平均智商仍然比控制组高出3.8～5.3分。不过这些孩子的学习成绩一直都没有很大的差别。

研究证明，参加这项计划时间最长的孩子，智商提高得最快。不过这种分析——显示出参与时间最长的孩子受益最大——难免存在自我选择的嫌疑，即使调查人员采取了一定的控制措施，也不能完全抵消自我选择的影响。而且需要注意的是，对于出生时体重极轻的孩子——不到2千克，这项计划在这些孩子3岁以后并没有取得多少效果。

另外一个非常重要的发现就是，和白人孩子（平均来说，白人孩子在这些干预措施针对的方面本身就具备更大的优势）相比，黑人孩子和拉美裔孩子从儿童早期教育计划中受益匪浅。同样，与中等阶层的孩子相比，穷人家的孩子受益更大。

总之，儿童早期教育干预计划对于处于不利地位的孩子和少数族裔的孩子是卓有成效的——只要研究人员尽心尽力地执行了这些计划。很多计划在结束时都使参与者的智力得到了大幅提高，但是在上小学期间这些效果就逐渐消失了。不过有一些证据显示，如果孩子们能进入高水平的小学念书，智力上的优势就不会被消磨殆尽。比提高智力更重要的是孩子们取得了更好的成就：接受特殊教育的孩子少了，留级的孩子也少了，参加标准化考试的成绩提高了，高中顺利毕业继续深造的人多了，违法犯纪的行为少了，收入提高了，靠救济生活的人也少了，而且有时这些变化之大还真是让人吃惊。

家访活动主要是为了改善父母养育子女的方式，不过一直没有多少研究对

家访活动的效果进行评估。一些投入了不少人力、物力和财力的养育干预计划，不仅改善了母亲的养育方式，还提高了孩子的情感和认知能力。成功的关键似乎就在于针对某些具体行为对家长进行培训。

苏珊·兰德里及其同事实施了一项卓有成效的养育干预计划，结果显示，这样的计划很可能具有极大的价值。参与该计划的大都是低收入的黑人和拉美裔美国女性，子女的年龄介于1～2岁之间。研究人员走进这些家庭，向这些母亲讲解如何对孩子采取更有益的回应模式，每次一个半小时，持续10～20次。他们还教妈妈们学会如何通过孩子发出的积极或消极的信号来识别孩子们的意图——当孩子没有服从母亲的要求时，如何以更加温暖、更加细腻的方式回应孩子们的行为；当进行互动和游戏时，如何吸引孩子们的注意力并激发和培养他们的兴趣，以及如何使用更丰富的语言来形容各种物体和动作。这项计划在很多方面对母亲们产生了显著的效果，包括对待孩子的方式更加温暖更加细腻，正确地回应孩子的意图，在各种活动中激发孩子的兴趣，在口头上对孩子进行鼓励。这些方面的效果值最高达到了1个标准差。

这项计划对孩子们的行为也产生了明显的效果。孩子们变得更愿意配合了，与母亲进行互动时更加投入，使用的词汇更加丰富，并且能够很好地运用语言描述他们周围发生的事情，并在一项图片词汇测试中得到了更高的分数。这些方面取得的最大效果值高达0.7个标准差。

虽然我们尚不清楚这样的养育干预计划长期效果究竟如何，然而按这样的结果来看，我们应该对此充满信心。而且这种计划的效益成本比也有可能会很高。

学龄儿童教育干预计划的实效

那么学龄儿童干预计划效果如何呢？对于那些从未参加过学前干预计划的孩子，我们应该做些什么呢？对于那些在幼儿园里曾经接受过一定学前教育干预的孩子，我们应该怎样做才能保持住他们已经取得的进步呢？

我得先说点儿坏消息——很遗憾，坏消息还不少。在第4章中我曾经讨论过为了提高孩子的学习成绩所付出的种种努力。钱本身并不能使所有孩子们在整体上取得任何进步，所以我们不能指望只用钱就可以缩小不同社会经济地位的学生以及少数族裔和白人学生之间的成绩差距。我们向穷孩子和少数族裔的孩子派发教育券，使他们有机会进入私立学校念书，但是没有什么证据可以证明这种做法能够真正解决问题。对于特许学校来说，情况也差不多。一部分特许学校比公立学校要好一些（实际上，你在后面会读到一个这样的例子），但即使是在已经开办了一段时间以后，特许学校在整体上似乎仍和公立学校的水平差不多，事实上，在开办之初还不如公

立学校。

那么有没有一些公立学校能够显著提高处于不利地位的少数族裔学生的学习成绩呢？两份不同的报告都坚持认为，至少有一部分学校使处于社会底层的少数族裔学生取得了极大的进步。一份报告出自一个保守派的机构，另一份报告则出自一个自由派的机构。

美国传统基金会是一个保守派的机构，该机构宣称已经找到了21所主要面向贫困学生的学校，而这些学校学生的成绩都超过了全美平均水平。他们认为这些学校之所以能够鹤立鸡群，是因为都拥有具有远见卓识的校长。这样的校长乐于抵制“对公共教育的盲目推崇”，能够辞退不合格的教师，强调向学生教授基础知识而不是难度过高的内容。

理查德·罗斯坦曾经是《纽约时报》国家教育专栏的专栏作家，他证明了美国传统基金会所说的情况并不属实。在他看来，这21所学校中只有6所是对生源完全没有选择权的社区学校，其余学校的情况是：（1）一部分是磁铁学校；（2）有些学校很多学生的家长都是“挣钱很少的”研究生；（3）有的学校注重基础知识，例如语音，所以在低年级时学生的分数确实很高，但到了高年级，学生的分数却直线下滑，这正是因为早期过于强调基础知识而忽略了对推理和解读能力的培养，而高年级的学生要想取得好成绩，以上两种能力必不可少；（4）有些学校家长必须提出申请孩子才能入学，因此也就存在自我选择的问题。

美国教育信托基金会是一个自由派的机构，其宣称找出了1 320所学校，这些学校里至少有一半的学生都是贫困的少数族裔学生，但是考试成绩全部位于所在州的前1/3的水平。不过罗斯坦认为美国教育基金会所说的情况也不属实。这1 320所学校的学生分数确实很高，但只是在某一年级里、某一科目上和某一学年里取得了高分。这样的成就只不过是统计上出现的意外情况而已。

道格拉斯·里夫斯宣称在密尔沃基找到了一些“90/90/90”学校，也就是说，在这些学校里，90%的学生都是贫困学生，90%的学生都是少数族裔，90%的学生都达到了“很高的学习水平”。上述发现也被罗斯坦揭穿了，所谓“很高的学习水平”实际上只是威斯康星州规定的基本的非熟练水平的分数而已。

还有人宣称，如果把穷困的少数族裔孩子送入家境富裕的孩子就读的学校，他们的成绩就会发生质的飞跃。情况真的如此吗？《纽约时报》2005年时在头版上（并分别于2007年和2008年在头版上以及杂志报道里再次刊登）宣称，在北卡罗来纳州罗利市的维克郡已经取得了这样的效果。这篇标题为“按家庭收入分配生源，罗利市学生成绩大幅提高”的文章写道：

在过去的10年里，维克郡的黑人学生及拉美裔学生在阅读和数学标准化考试中的成绩迈上了好几个台阶，这已经引起了美国教育专家们的关注。

维克郡（包括郡政府所在地罗利市及附近的郊区）的官员和家长们都表示，孩子们取得这么大的进步，要归功于这一学区一直同心协力地运用经济手段来促进学校的融合。

文章接着介绍说，从2000年开始，官员们就开始按学生的家庭收入将学区的生源分配到不同的学校，目的是为了将学校中低收入学生的比例控制在40%以下。结果如何呢？

十几年前，维克郡三到八年级的黑人学生中，只有40%在全州统考中的成绩达到年级平均水平。后来，这一比例已经上升到了80%。拉美裔学生的进步幅度与之类似。

总体上看，这些年级中有91%的学生达到了年级水平，而10年前这一比例是79%。

不幸的是，并没有证据显示罗利市的学校比全州其他学校更有利于少数族裔学生的学习。实际上，在2004—2005学年度，维克郡的学校中黑人与白人学生的差距大于全州整体水平。不仅如此，几乎可以肯定的是，维克郡的学校中少数族裔学生取得的进步主要是因为全州统考的标准降低了，学生们更容易拿到熟练等级的分数。按照新标准注，95%的白人学生都可以达到熟练的等级。这就是说，这个标准太低了，根本不可能完全显示出黑人学生与白人学生的真实差距。

不过，将维克郡与全州整体水平进行比较之后，还是发现了一个令人欢欣鼓舞的现象。那就是，家庭收入低的学生与家庭收入高的学生在同一所学校里学习，对彼此都没有什么坏处——或许还能使双方取得一些社交方面的收获。不过我们还需要了解更多的情况——例如，全州及罗利市的黑人与白人的社会经济地位水平——才能得出这样的结论。不过有充分的证据显示，和在主要由黑人或拉美裔学生构成的班级里学习相比，黑人及拉美裔学生在种族融合的班级里学习效果更好。

那么，对学校的管理和学生的课程安排进行自上而下的改革会取得什么样的效果呢？从第4章中我们了解到，有很多这样的全面的学校干预措施，对课程及教学策略进行了全面的改革。不过这样的干预措施大都没能显著提高学生的整体学习成绩，所以我们也就对这样的措施能否有效地缩小学习差距心存怀疑。

教师的资质及高学历也不一定能提高学生的学习成绩。执教的经验倒是很

关键，并且与社会阶层较高的学生和白人学生相比，教师的经验很可能对社会阶层较低的学生和少数族裔的学生更重要。教师素质对学生的成绩也起着重要的作用。同样，很可能教师的素质对穷孩子和少数族裔的孩子更重要。我们当然还记得第4章中提到的A女士。她对曾经教过的一年级的穷孩子产生了巨大的影响。另外，对于那些穷孩子和在学前班时就很令人头疼的孩子来说，他们在一年级时得到的教学支持和情感支持会对他们的学习成绩产生巨大的影响。

我们还知道，学生在人数较少的班级里学习时学习成绩更好，这对黑人学生的效果要大于白人学生——分别是0.33个标准差和0.25个标准差——对穷孩子的效果要大于中等阶层的孩子。

不过学龄儿童教育干预在基础教育（从学前班到高中三年级）方面倒是有一些成功的故事可以分享。穷困的少数族裔孩子参加的两个教育项目（一个是数学项目，另一个是阅读项目），现在已经取得了不少成效。

数学训练项目叫作种子计划，是一项课外补充项目，专门聘用并培训数学家、科学家以及工程师，让他们给贫困的少数族裔学生上数学课。教师们采用苏格拉底的问答法，逐步引入抽象的数学概念，鼓励学生积极参与课堂教学——采用对话、辩论和集体回答的形式。全职教师会旁听全部的课程。这项计划并不是要取代平时的数学课，而是一种课外学习。一项在达拉斯进行的研究对种子计划的效果进行了评估。该研究选取了244名参加了种子计划的四年级学生，然后在这些学生所在的学校内选取了244名没有参加该计划的四年级学生，又从其他学校里选取了244名和这些学生成绩接近的四年级学生组成控制组。研究人员将这3组学生在加州成就测验中的得分进行比较后发现，参加种子计划的学生其成绩比控制组的学生高出0.37个标准差——这算得上是很大的差距了。不过，参加了种子计划的学生的成绩，只比同一学校内未参加该计划的学生高0.19个标准差。未参加种子计划的两个组都可以作为控制组，但是选择哪一组进行对比会有很大的差别。这项计划的成本不算太高，能够取得0.37个标准差的效果值当然是值得投入的，不过如果效果值只有0.19个标准差，或许就很难说了。

在一些面向穷孩子和少数族裔孩子的阅读项目中，最有效果的要属阅读恢复计划了。这项计划是俄亥俄州州立大学的研究人员开发的，旨在辅导学习成绩较差的一年级小学生进行阅读训练。该计划对孩子提供一对一的辅导活动，每次30分钟，每天一次，持续12~20周。孩子们读一些他们听过的故事，重读一遍前一天读过的故事，再自己写一些故事。研究人员随机选取了一些孩子，对计划的效果进行了研究。他们发现，在绝大多数反映阅读熟练度的指标上，这一计划的效果值是从0.57到0.72个标准差不等。虽然这些效果随着时间渐渐流失了，但直到孩子们三年级时仍有大约0.2个标准差的效果。对于这项计划中的西班牙语项目进行评估的独立研究发

现，不同指标的效果值是从1到1.7个标准差不等。不过我们仍然需要对这项计划进行更多的重复测试，才能完全相信它的效果。

全面学校干预措施和特许学校对学生的成绩没有太大的作用，但这种说法至少存在一个极其重要的例外：KIPP（Knowledge Is Power Program，知识就是力量项目）。这个项目的创始人迈克尔·范伯格和戴维·列文曾经是休斯敦的小学教师，两人教龄相差4年，在1994年携手开创了这一被寄予厚望的教育项目。他们根据贫困学生尤其是少数族裔学生的需求制订了这项计划。他们坦承当时也是一边上课一边摸索。不过他们得到了一位经验非常丰富的老师的帮助——哈丽雅特·鲍尔——她在休斯敦一个种族隔离的社区里长大，但她教出来的学生个个都行为端正、成绩优异。

范伯格和列文建立的学校最初主要面向中学生，上课时间是从早上7点半到下午5点（没错，每天9.5个小时），暑假必须再上三周课，每隔一周的周六上半天课。他们会定期进行家访，要求学生做到心地善良、行为端正。校长有权聘用或辞退教师，教师之间要通力合作。学校还建立了奖惩机制来监督学生的行为和学习成绩。而且学校还有大量额外的接触时间，使得学生们可以接触到中上阶层的学生在家里或昂贵的公立或私立学校里进行的活动——体育运动、参观博物馆、舞蹈、艺术、乐器、剧院和摄影。最早建立的两家学校——在休斯敦和布朗克斯——在这两个地区所有的学校里成绩排名第一。从2001年开始，KIPP得到了来自杜丽斯和唐纳德·费舍尔（盖普品牌的创始人）的资助，开始在各地建立连锁学校。大部分KIPP学校都是特许学校，同时几乎全部都是KIPP基金会的特许经营学校。还有两项计划与KIPP极为相似——成就第一计划和北极星计划，但并没有得到更多的研究。

KIPP的学生几乎都是家庭极其贫困的学生，超过80%的学生都符合申请联邦免费或补贴午餐的条件。大部分学生都是非洲裔和拉美裔。KIPP坚信“进入KIPP的五年级学生在全国统考中平均位于最底部的1/3（百分位排名第28位），但当他们八年级参加阅读和数学的常规参照性考试时，得分比全美3/4的学生都要高（百分位排名第74位）”。

不过也有一些KIPP学校里的学生，成绩并没有取得多大提高（值得称赞的是，这是KIPP学校自己主动承认的），这些KIPP学校所宣称的成绩大都是学生们在KIPP的老师自己设置的考试中取得的，而不是经独立研究评估得到的。21世纪初的几年，美国斯坦福国际研究所对位于旧金山湾区的五所KIPP学校进行了独立研究，下面我详细地介绍一下研究的情况。

在旧金山湾区的5所KIPP学校中，72%的学生家境贫寒，75%的学生是非洲裔和拉美裔。研究中的每所学校都和两所学校进行比较，这两所学校中学生家庭的经济状况和少数族裔学生的比例都和KIPP学校很接近。在

2002—2003年度，KIPP学校只有一个五年级的班级，以后逐年增加一个班级。

学校的校长要对KIPP的模式进行一年的研究，才可以开办自己的特许经营学校。不过KIPP模式并没有规定特别的教学方法或课程，所以教师们不用学习一个新的教学方法体系，这和大多数的全面学校干预措施不同。在进行研究的时候，旧金山湾区所有KIPP的教师中，大约有一半都来自“为美国而教”项目，他们在加入KIPP时的平均教龄是两年。

KIPP学校的学生也有可能考试不及格。校长们认为这样做是非常必要的，因为学校不能“让其他孩子看到不努力的孩子的成绩也能得到提高”。教师、学生和家长都要记住，有一些行为他们是必须要做到的。KIPP的信条是：“如果有问题，我们要找到解决办法。如果有更好的方法，我们一定要找出来。如果我们需要帮助，我们就向他人求助。如果队员需要帮助，我们就要伸出援手。”一些典型的KIPP的口号有“努力学习，礼貌待人”，“我们每个人都要学习”，以及“KIPP的人即使在没有别人在场的时候，也只做对的事情”。

学生们每周都会得到“支票”，支票的数额和记分系统挂钩，记分系统会根据他们的行为和学习表现增加或减去一定的分数。学生们可以用支票在KIPP商店里买东西，包括零食和学习用品，也可以用来购买校外考察的参与权。学生们如果行为违反规定或没有完成任务，就要受到单独坐在“长凳”上的惩罚。在一所学校里，学生们甚至只有连续三天支票上没有被扣去任何数额时，才能离开“长凳”。教师们向美国斯坦福国际研究所的研究人员再三强调，学生的所有行为都会得到奖励或惩罚。不过教师们坚持认为，学生们并不是因为害怕被惩罚才约束自己的行为。

“在我们的学校里，从来没有一个孩子在背后褒贬老师，孩子们也从来不打架。我认为这不全是由于纪律的约束，而是从一开始就让孩子们感受到自己背负的期望。即使是最小的细节我们也不会放过……这是因为孩子们相信这是一个特别的地方，我们也一直这样教育他们。我相信他们不会因为害怕“长凳”惩罚就不去取笑他人，而是在KIPP学校里他们不会这样做。KIPP就是一所学生不会遭到取笑的学校。孩子们在这里感到很安全，所以学习效果更好。”

另一位老师说：“在这所学校里，做一个聪明的学生不会有任何压力。这是大多数市中心的公立学校所缺少的……在我原来任教的学校里，我面对的是一直以来都缺乏帮助的孩子，我发现学校不能完全满足孩子们的需要……我去参观了一所附近的KIPP学校，感觉就像在沙漠里发现了绿洲一样。”

从学生们的评价来看，他们完全清楚KIPP学校与他们曾经上过的其他学校之间的差异。“每个人都努力学习”，“其他的学校没有这里要求高，我发现这里处处都充满了挑战”，“现在我明白了要想上大学就必须用功读书”。在这些学校中，学生们都说，和他们以前就读的学校相比，这里的学生很少打架，实际上几乎没有谁打架：“在这所学校里，好像有什么东西在阻止我和别人打架。”

当然，学校对于教师的要求很高。教师们从早上7点15分到下午5点15分要一直待在学校里，通常实际待的时间还会更长一些。除了备课以外，他们要上课，监督学生们上自习，组织各种课外活动，例如参观动物园和博物馆，对学生进行辅导，还要在晚上给学生家长打电话沟通学生的表现。暑假里他们还要工作几周，隔周周六还要上班。这样一来，他们自然身心俱疲。绝大部分KIPP的教师都表示，这样下去，他们恐怕只能坚持几年。

所有这些努力会有什么样的成效呢？在旧金山湾区KIPP学校里念书的孩子，成绩远远高于根据他们的家庭状况和种族预测出来的成绩。学生们分别参加了春季和秋季的斯坦福成就测试（SAT10）中的阅读、语言艺术和数学测试，这些测试都是全美统考。对于五年级的学生来说，秋季考试的分数是他们刚刚加入KIPP时得到的分数。我们以百分位排名在第50位或更前的学生的比例，来衡量孩子们从秋季到春季取得的进步。秋季五年级的语言艺术考试中，有25%的孩子百分位排名在第50位或更靠前（所有参加考试的四所学校的平均成绩），稍稍高于我们根据学校的学生构成预测出的比例。到了春季考试时，这一比例上升到44%，远远高于根据学校的学生构成预测出的比例。秋季五年级的数学考试中，这一比例是37%，春季时上升到了65%。六年级的学生同样取得了显著的进步。如此大的进步显示出，在仅仅就读于KIPP学校一年之后，占学生总数绝大多数的少数族裔穷孩子，在标准化考试中的得分就已经接近或者超过了全美的平均水平。由于这些学生年初时的分数并不算太高，我们就可以排除春季时取得的成绩完全是自我选择的结果这一可能。

春季，所有孩子（全部五所学校）都参加了法律规定的加州成就测验。英语语言艺术的考试中，所有KIPP五年级学生中有43%达到了熟练及以上的水平，而在一所学生构成与KIPP接近的学校里，这一比例是19%。数学考试中，KIPP五年级学生中有55%达到了熟练及以上的水平，在与之对比的学校里这一比例是20%。六年级学生在加州成就测验中的成绩与五年级非常相似。同样，他们也取得了惊人的进步（基准线是：KIPP学校和与之对比的学校在年初加州成就测验中的成绩接近）。

在介绍SRI的研究时，我违背了我自己定下的绝对准则。在这项研究中，在KIPP学校念书的学生并不是随机选取的。不过我们可以这样做：有时会有一些孩子等待进入KIPP学校念书，KIPP就会采取抽签的形式来录取学

生。没有抽中的学生被编入随机分配的控制组，并且参加KIPP学生参加的考试。其实KIPP也担心他们录取的都是成绩已经很不错的学生。令人称赞的是，有的KIPP学校已经采取了一些措施来保证尽可能多地招收处于不利地位的少数族裔学生，因为这才是他们的目标生源。

不过这里面依旧存在严重的自我选择问题，因为决定孩子是否要进入KIPP学校学习的是家长，而不是研究人员。这就意味着，KIPP学生所取得的成就中，有一部分或许并不是学校努力的结果，而是在家长或其他优势的作用下取得的。但问题是，自我选择能使学生们取得这样大的进步吗？

在回答这个问题之前，我们至少可以基本肯定的是，有些学生由于成绩不理想而退学，并不是KIPP学校取得辉煌成绩的主要原因。在SRI研究的5所KIPP学校中，每年只有9%的学生中途退学，其中一部分是因为家长担心影响孩子的发展。学校会劝退一些学生，不过实际上没有一个学生是被学校开除的。退学的学生数量极其有限，所以学校造成的自我选择似乎并不是孩子们取得高分的主要原因。

总之，选择了KIPP学校的家长一定比类似阶层和家庭条件的家长更具远见，这一点我并不怀疑，不过我不相信KIPP取得的成绩都应该归因于自我选择。尽管如此，在社会将大量资金投入到KIPP学校之前，我们还需要对KIPP进行一些完完全全的随机研究。

我坚信，这样随机选取研究对象的研究会证明穷困的少数族裔学生（至少这些孩子的父母重视孩子的发展才把他们送进KIPP学校念书）的学习成绩完全能够达到中等阶层白人学生的成绩水平。下一步[注]就是要找出，那些不太重视子女教育的家长是不是也能从KIPP这种类型的计划中受益。

少数族裔学生的数学成绩也可以很优秀

你或许看过电影《为人师表》。这部电影讲述的是数学教师詹姆·艾斯卡兰德的故事，他使东洛杉矶南美移民的子女——那些通常连高中都无法毕业的学生——通过了大学预修课程的微积分考试，并且考试通过率超过了在贝弗利山高中就读的富家子弟，从而也就超过了全美大多数英才高中的学生。但电影里的故事是真实的吗？

关于艾斯卡兰德的举措，人们褒贬不一。最重要的是，电影中的故事绝非杜撰，而是确有其事。

遗憾的是，真实情况并不像电影中那样。在现实生活中，艾斯卡兰德并没有向信任他的学生们承诺他要把他们培养成为数学高手。他在直属初中开

发了数学课程，参加这些数学课程的学生都以优异的成绩升入了他任教的高中。他确保这些学生在成为他的学生之前，就可以受到很好的高中数学教育。他做到了，但遭到了他任教以来就受其领导的第一位校长的强烈反对。后来学校换了一位支持他的校长（这位校长禁止平均成绩低于C的学生加入体育团体），事情才逐渐好转。后来艾斯卡兰德又和教师工会发生了激烈的冲突，因为他所带的班级人数大大超出了工会规定的学生人数。因为他找不到足够多的好老师来开设更多的班级，所以班级的规模越来越大。

后来另一位不怎么支持他的校长取代了支持他的校长。艾斯卡兰德最后因为班级规模问题和其他问题离开了学校。在他走后，他所开设的数学课程也没有以前那么成功了，不过他曾任教的高中的大学预修数学成绩，还是远远好于大多数同类学校。

艾斯卡兰德所取得的成绩具有非常重要的意义。它以铁一般的事实证明，处于不利地位的少数族裔学生也可以取得远远超出全国平均水平的数学成绩。

社会心理学家提出的简单且效果显著的干预措施

我的一些社会心理学家同行最近登上了教育的舞台，他们带来了一些既容易操作又非常划算的干预措施。

很多美国人都相信，一个人的能力在出生时就决定好了：要么你生来就有数学天赋，要么生来就没有。其他人则认为能力完全可以通过锻炼得到提高：如果你努力的话，你在某一方面的能力就一定比你不努力时的能力强。卡罗尔·德韦克及其同事在一所主要由少数族裔学生构成的中学里开展了一项调查，调查学生们对于能力的态度。他们通过学生们对一些问题的回答来了解他们的看法，例如，“每个人的智力都是由遗传决定的，没有什么方法可以改变人的智力”以及“你一定可以变得比现在更聪明”。调查的结果正如人们所料，认为付出努力就可以提高能力的学生，其学习成绩好于那些认为能力完全取决于基因的学生。

随后，德韦克及其同事努力使一群穷困的少数族裔初中学生相信，智力绝不是一成不变的，付出努力就可以提高智商。这项干预措施的要点在于，学习可以使大脑长出更多的神经连接，进而改变智力，所以学生们完全可以掌控这一改变的过程。德韦克介绍说，当听到智力在很大程度上可以由人们自己来控制的时候，一些平时非常难以管教的初中男生竟然掉下了眼泪。据老师们反映，接触到这一干预计划的孩子学习更加用功了，并且他们的分数也高于控制组中的孩子。和倾向于相信付出努力就可以提高智力的孩子相比，这项干预措施对那些最初相信智力完全由基因决定的孩子更

有效果。

乔舒亚·阿伦森及其同事进行了类似的实验，得到的结果也很激动人心。其中一项研究是在得克萨斯州一些穷困的少数族裔学生中进行的。他们对这些学生进行了强化的干预措施，取得了惊人的效果。

在这项研究中，每一名学生在初中一年级时都被配备了一位大学生辅导员。这些辅导员会和学生们谈论关于如何适应学校的一系列问题。控制组的辅导员会和孩子们讨论毒品的问题，并鼓励他们远离毒品。干预组的辅导员会告诉学生智力是可以改变的，并指导他们怎样做才能使大脑在一生中不断形成新的连接。干预组的每个学生都能看到一个网站，这个网站用动画的形式向学生展示大脑的构造，包括神经元和神经细胞树突的图片，并配有语音介绍，向孩子们解释当新问题被解决时大脑是如何长出新的连接的。辅导员们还帮助孩子们设计了一个网页，孩子们用自己的语言和自己亲手画的图，在网页上展示了他们从辅导员那里接收到的信息。这一切都强化了辅导员向他们传递的信息。

这项干预措施的效果非常显著。在得克萨斯州学业能力评估的数学测试中，干预组中男孩的成绩比控制组中男孩的成绩高出0.64个标准差。女孩往往担心由于性别原因学不好数学，而干预组中的女孩的成绩高出控制组1.13个标准差。在阅读测试中，干预组的平均成绩比控制组高出0.52个标准差。

达芙娜·沃耶斯曼及其同事精心设计了一项专门针对穷困少数族裔初中生的干预措施。他们开办了几期活动，目的是要让孩子们仔细考虑一下他们究竟想要什么样的未来，在通往未来的过程中他们会遇到哪些困难，他们如何处理这些困难，他们有哪些朋友最能帮助他们解决这些困难。他们同时还创办了几期活动，把孩子们分成小组，讨论如何解决日常生活中出现的问题，社交上遇到的困难，学习上的问题，以及如何为顺利从高中毕业做好准备。这项干预措施使参与了这些活动的孩子们的成绩平均分提高了0.23个标准差，在标准化考试的成绩的效果就更好一些，有0.36个标准差，留级的可能性也显著减少，降到0.6个标准差。

即使是一些很小的干预措施，在大学里也同样会引起一些改变。很多大学校园里的学生都担心是否能够被学校接纳，自己能否融入学校生活，少数族裔的大学生在这个问题上尤为苦恼。如果他们交不到朋友，因为校园里没有那么多少数族裔学生，或者他们没法自在地和白人学生相处，他们或许就会怀疑自己是否真的适合大学生活。随着学习的深入，少数族裔的学生会逐渐丧失动力，成绩平均分也越来越低，这在大学校园里是很常见的事情。

社会心理学家格雷戈里·沃尔顿和杰弗里·科恩认为，每个学生都会担心自己能否被学校接纳，这与种族无关，目前的状况在将来很可能会有所改观。如果少数族裔学生能够认识到这一点，学习成绩每况愈下的情况就可以被消灭在萌芽状态。研究人员对一所名牌私立大学里的黑人学生采取了一项适当的干预措施。他们邀请黑人和白人学生在大一即将结束时参加一项心理研究。实验人员的目的是要让干预组的学生相信，担心自己能否被学校和同学接纳是很常见的问题，随着他们建立起更多的友谊，这种忧虑就会慢慢消失。研究人员希望这样能够帮助黑人学生认识到，不应该从种族的角度来理解自己面临的社交困难（“我想我们这类人真的不太适合这种地方吧”），而是要相信大家都面临这样的问题（“我想每个人都会遇到这样的问题”）。研究人员相信，只要他们认识到这个问题的普遍性——以及可能的对策——就不会再担心归属感的问题，而是会将精力集中到学习上来。

言归正传，沃尔顿和科恩让干预组中的学生每人写一篇文章，内容围绕着他们在未来是否有可能改善自己的社会地位，并且要求他们在摄像机面前发表一个简短的演讲。研究人员告诉这些学生，他们的演讲会被放给学校的新生们听，“以便让他们了解大学生活究竟是什么样的”。一周以后，研究人员调查了这些学生的学习成绩，以及下一学期的成绩平均分。

这项干预措对黑人学生起到了巨大的积极作用，不过对白人学生却没有什么成效。根据报告，参与干预措施之后的那段时间，黑人学生学习的时间变多了，和教授们的接触也比以前频繁，参加复习课和学习小组会议的次数也增加了。接下来的一个学期，干预组中黑人学生的成绩也反映了这些情况：他们的分数比控制组中的黑人学生整整高出1个标准差。

发挥大学消除智力差距的作用

实际上，大学会对黑人学生的智力产生巨大的影响，这种影响比对白人学生的影响要大得多。

黑人与白人的智力差距在高中时开始逐渐拉大。一些遗传决定论者认为，基因在发展的过程中会逐渐显现出它的威力，所以黑人与白人的智力差距在高中时才会越来越大。随着教育水平的提升，黑人越来越被甩在后面。例如，赫恩斯坦和莫瑞就坚持认为，高中以后的教育不可能缩小种族之间的智力差距。

显示黑人与白人的智力差距在高中时逐渐增大的数据，来自全美青少年纵向调查，这个机构管理着武装部队职业资格测试。武装部队职业资格测试是军队对未来的士兵进行测试时采用的一系列能力测试的一部分，由于与智力测试高度相关，所以也可以将其视为是智力测试。所有全美青少年研

究调查的参加者，都在14～21岁的不同时间里接受了武装部队职业资格测试。

不用说，正如赫恩斯坦和莫瑞所说，在高中时黑人的能力提高得比白人慢，两者之间差距增大的速度着实令人担忧。刚进入高中时，黑人在武装部队职业资格测试中的成绩比白人低了不到0.6个标准差，但当高中毕业时，黑人在武装部队职业资格测试中的成绩比白人低了将近1个标准差。

心理学家乔尔·迈尔森及其同事开始着手调查黑人在大学里是否也会出现这种情况。如果按照能力差距随着时间会越来越明显的说法，黑人学生的智商在大学里的提高值就会比白人学生的小，所以两者间的差距应该比高中时还要大。

但迈尔森及其同事得到的结果恰恰相反。最终从高中毕业的黑人，比最终从高中毕业的白人的智商低1个标准差。但是白人学生在大学期间智力增长极少，而黑人学生的智力得到了极大的提高，到大学毕业时，黑人学生的平均智商与白人学生的平均智商的差距仅稍稍大于0.4个标准差。大学教育给黑人学生和白人学生带来的附加值有着巨大的差别。

为什么黑人在大学期间能够取得如此大的进步呢？恐怕把这个问题换成“为什么他们在高中时的进步那么小呢”更合适。最容易想到的答案就是他们读的高中比白人差。另一个答案就是，黑人学生在高中时比在大学时更难面对被同伴指责为“模仿白人”的压力（如果他们在大学也面临这样的压力的话）。

第三个可能的答案来自有关刻板印象威胁的研究。这项研究显示，黑人学生的测试表现和积极性有着很大的差别，这取决于他们面临的社会环境。例如，斯蒂尔和阿伦森的实验研究里有记录显示，当采用一种不会让黑人学生感到威胁的方法进行测试时——也就是这些学生没有明显感觉到这些测试测量的是他们的智力时，他们的表现就会好很多。这是因为如果他们意识到这是智力测试，他们就会变得紧张，生怕他们的表现恰恰印证了别人心目中的黑人智商低的印象。这会影响他们在测试中的表现，而且黑人学生比白人学生更有可能采用消极的方式来面对这种评估行为带来的不适，例如回避挑战，对自己的学习成绩漠不关心，这些都是黑人初高中时面对刻板印象威胁的典型反应方式。一项跟踪调查黑人学生高中生涯的研究发现，整个高中阶段，黑人男学生花在学业上的时间急剧减少，所以到了十二年级的时候，他们的自我价值感与学习成绩就一点儿关系都没有了。那些接受了他人对黑人的刻板印象的学生，在这一点上表现得尤为突出。另外一项长期研究发现，黑人学生越是担心别人对黑人持有的消极印象或许是真的，他们就越不会用功学习。所以那些将刻板印象威胁最大化的情形，可能在高中校园里尤其普遍。

不过，为什么黑人会在大学里取得更大的进步，我们不得而知。我们能够确信的是，大学教育能够显著缩小智力上的差距。遗传决定论者认为，随着时间的流逝，智力上的遗传劣势就会发挥越来越明显的作用，所以随着年龄的增长，智力的差距会越来越大。而可靠证据显示，实际上智力差距在大学期间会显著缩小。

总结

那么，现在关于针对穷孩子和少数族裔孩子的干预措施，我们都了解了哪些内容呢？有些内容令人感到非常意外。也许我们学到的最重要的一点就是，只有实验才能证明哪些措施有效，而哪些措施无效。

有些早期儿童教育计划看起来科学合理，但实际上没有多大——或持久——的效果。先行教育计划是个很明智的计划，本应该取得很好的效果——在早期也确实取得了很好的效果。但是当孩子们回到贫穷的家里，简陋的社区，进入条件很差的学校，孩子们在智力上取得的进步就会慢慢流逝，学习成绩也没有太大的起色。不过，一些学前干预计划最初确实极大地提高了孩子们的智力，如果孩子们可以进入更好的公立学校念书，这些取得的显著效果就不会消失。即使孩子们没有进入好学校念书，那些效果最好的计划还是可以对学习成绩产生巨大的作用，并且对社会也大有裨益，例如减少犯罪和靠救济生活的人的数量。大体上，学前干预计划对穷孩子和少数族裔孩子所起的效果，大于对家境较好的白人孩子所起的作用。

小学和初中的干预措施也存在同样的情况。我已经介绍了不少对学生学习成绩没有太大作用的干预措施，所以也不可能消除种族之间和社会经济地位之间的差距。真正能够带来改变的是：教师个体的素质、对教师发放奖金（有这种可能，不过还需要做很多工作才能找出最有效最实用的奖励计划），以及班级规模（对黑人学生的效果比白人学生大）等因素。

我们在第4章中讨论过的一些教学方法花费不多，但效果却相当好，或许这些方法有助于缩小学业成绩上的差距。这些教学方法包括，使用电脑程序来教授数学和写作（这实际上也许并不比标准教学方法更昂贵）以及“合作学习”（几乎不会增加任何成本）。即使这些方法都被证明不能缩小差距，但它们也确实能从整体上提高孩子们的学习成绩，所以应该被用来教育所有类型的孩子。即使涨潮将所有小船托起来的高度都是一样的，这样的涨潮也是受欢迎的。

我们知道有一个数学项目——种子计划，和一个阅读项目——阅读恢复计划，在少数族裔孩子身上取得了极大的效果。尤其是阅读恢复计划，成本不高，但取得的效果却很大。

我们已经尝试过大量的全面学校干预计划，但结果大都令人失望。尽管如此，还是存在一个极其重要的例外——KIPP。就像詹姆·艾斯卡兰德一样，KIPP学校的教师能够使贫困的少数族裔小学生的数学成绩，提高到中等阶层白人学生的一般水平。即使是较晚进入到KIPP学习也可以做到这一点：五年级时加入KIPP的学生取得了惊人的进步。不过我们还需要确定年龄较小的孩子加入KIPP能获得多大的提高，让我们期待这一天早日来到。

为消除差距买单

不过，我们能承担得起有效干预措施的代价吗?或许这样问更好：我们能承担得起不采用这些有效干预措施的代价吗?许多经济学家都曾经对最成功的学前干预措施的收益成本比进行过计算。荣获诺贝尔奖的经济学家詹姆斯·赫克曼就对佩里学前教育研究计划的回报进行过评估——包括由此节省的课外辅导的费用，留级的孩子重读的费用，犯罪和福利的支出，以及由此给参与者带来的未来收入的增加——收益与成本的比是8：1。这相当于年投资收益率达到17%。而这计算的仅仅是金钱方面的收益，并不包括参与者、参与者的家庭和邻居们的生活质量的提高。佩里项目的初期成本很高——按照2007年的美元价值计算，估计每个学生的成本在12 000～16 000美元。这比先行教育计划的成本高多了（不过以后类似佩里项目的项目成本就会有所下降，因为不再需要支付研究费用，而研究费用占到了佩里项目所有成本的很大一部分），但是效果也要好得多。初学者计划也是如此。这项计划的收益成本比估计是3.78：1。即使只计算与纳税人有关的收益，包括由此省下的教育、福利和刑事审判的支出，这项最成功的学前干预计划最终也能收回成本。

如果让家庭最贫困的1/3的孩子从出生到进入学前班之前一直参与佩里计划或初学者计划，总共的费用是多少呢？在美国大概有700万这样的儿童，佩里计划或初学者计划花在每个孩子身上的费用每年约为15 000美元。这样计算下来，总共的费用就是1 050亿美元。不过先不用太震惊，因为我接下来要告诉大家，现在国家对各种学前干预计划的投入也已经达到200亿美元了。这还没有包括一部分私人在教育上的投资，母亲们在计划实施期间及计划结束后的几年间增加的收入，以及孩子们和社会最终得到的收益。不仅如此，我们也并没有对成本最高的项目进行充分的研究——有可能花在每个孩子身上的钱远远少于15 000美元，就足以使认知和社会职能取得极大的进步。最后请注意，学前项目的收益和需要是成比例的。母亲的智商和社会经济地位越低，从学前项目中获得的收益就越大。所以即使我们只向最贫困的1/6或1/12的5岁以下儿童提供大力度的干预计划，回报也是非常大的。

请注意，2001年以后美国对最富裕的1%的家庭实行减税，光是这一项就

会使财政部2009年的收入减少940亿美元。我们可以将这个数字作为衡量儿童早期教育总开支成本大小的标准。

那么效果极为明显的KIPP呢？KIPP学校的支出实际上比一般的公立学校多不了多少。（一些KIPP学校比它们所在地区的公立学校支出还少。）不过这都是因为KIPP学校的教师大多年轻、有理想，所以心甘情愿地为学校付出，而他们的工资仅仅略高于公立学校里同等水平的教师。但KIPP的老师无法一直都保持这样的工作强度，这一点儿也不奇怪，工会已经开始就这些问题与KIPP学校进行交涉了。

如果向KIPP学校的老师支付和公立学校教师相当的小时工资——这样做不仅是出于公平的考虑，而且是十分必要的，只有这样才能为学校找到足够的教师——学校的支出会是多少呢？KIPP学校的老师与学生接触的时间比别的学校多出60%，大部分的教育成本都是实物产业成本，行政管理和维修成本，以及负债利息，这些费用都与KIPP采用的教学方法无关。2005年美国公立学校学生的平均教育支出是8 000美元，其中1/3都用于支付教师的工资。如果我们假设教师的工资增加了60%，4 000万名5～14岁的美国儿童中有1/3参加类似KIPP这样的计划，那么成本就会增加350亿美元。同样，由此省下的儿童托管费用以及孩子母亲增加的收入能够抵消这些成本的一大部分。虽然现在还不能计算出加入这样的学校的孩子在一生中能够获得的经济收益，但至少能够用来支付一大部分这些多出来的成本。

要说明的是，我并不赞成现在实施任何早期干预计划。类似佩里计划、密尔沃基计划或初学者计划的早期干预计划，在全美范围内推广是否有效和可行，我们还需要进行大量的研究才能够最终确定。类似KIPP的计划也是如此。对于并没有努力争取使子女进入KIPP念书的家长，他们的子女能够在KIPP项目中取得多大的进步，这才是我们要继续关注的。

有大量证据能够证明，我们确实有可能显著缩小人们智力和成就之间的差距，并且我们清楚，有效的干预计划的成本至少是可以接受的。既然如此，我们就必须进行必要的研究，找出最有效的干预计划，否则就是不负责任。

最后，如果我们想使穷人变聪明，或许先让他们变得富裕一些才是更好的选择。斯堪的纳维亚国家在收入分配上比美国更具平均主义倾向，该国最穷的孩子和最富的孩子之间的差距也反映了这种相对的均等。诚实守法地守着一份符合社会价值观的工作就应该能够养家糊口。提高最低工资（即使最近增加后的最低工资也只相当于40年前的73%），劳动所得税扣抵制以及儿童税收抵免或许能够使这个问题得到部分解决。

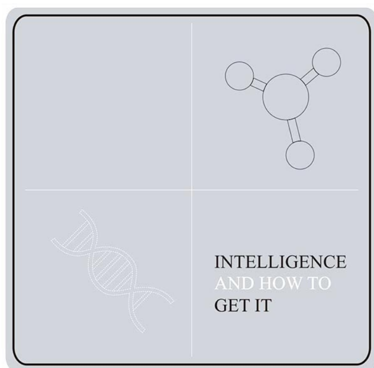
提高穷人的生产力，降低犯罪率和靠救济生活的比例，至少可以弥补这些减免政策的部分经济成本——甚至有可能不止是成本。帮助别人或许真的就是在帮助我们自己。

---

1. 2005—2006年的数学成绩更能说明问题，因为熟练一级的分数档比2004—2005年降低了许多。正如很多州的情况一样，北卡罗来纳州熟练一级的分数档每年都会进行调整，而调整的依据并不是学生的实际成绩水平，而是出于政治的需要。2005—2006年，数学成绩熟练一级的分数档再度提高时，结果也是一样的，维克郡的学校中黑人与白人学生的差距大于全州整体水平。纽约时报那篇报道的作者应该从源头处获取数字，而不是依赖教育者提供的数据和解释，这些教育者也是完全参考维克郡的区报告才获得这些数据的。

2. 还有另外两个设计与KIPP极为类似的项目，一个叫作成就第一，主要在康涅狄格州纽黑文市开班；另外一个叫作北极星，主要在新泽西州的纽瓦克市开班。这两个项目宣称，他们的学生类似于KIPP学校的学生，这些学生也同样能取得极大的进步。不过据我所知，这两个项目都没有得到很好的评估。

第8章
亚洲人智商更高还是更努力?



天啊，只有亚洲人才能得到这样的分数！

——一名就读于硅谷一所欧美合办高中的高三学生听到她的同学得到了极高的**SAT**分数后的反应

是故无冥冥之志者，无昭昭之明；无惛惛之事者，无赫赫之功。

——荀子《劝学》

下面这些统计数字应该能够引起具有欧洲血统的人的注意。

- 1966年，在美国高中学习的华裔高三学生参加SAT考试的人数比欧裔学生高出67%。尽管并不是最优秀的学生，但平均来说，这些华裔学生的分数和欧裔学生的分数非常接近。

- 1980年，“66级”的华裔学生已经32岁了，他们中成为专业人士、管理人员或技术人员的比例比欧裔学生高出62%。

- 20世纪80年代末，在加利福尼亚州橘郡的加登格罗夫，东南亚“船

民”的子女只占当地全部人口的20%。然而在当地的14所高中里，有12所学校都是由这些孩子作为毕业生代表在毕业典礼上致辞的。

- 1999年，第三次国际数学和科学研究的结果显示，美国八年级学生的数学成绩比日本、韩国、中国大陆和港台地区、新加坡的同年级学生低0.75~1个标准差，科学成绩低0.33~0.5个标准差。

- 虽然亚裔美国人只占美国总人口的2%，但是2008年西屋科学大奖赛的5位获奖者全部是亚裔美国人。

- 哈佛大学现在有20%的学生都是亚洲或亚裔学生，加州大学伯克利分校有45%的学生是亚洲或亚裔学生。

所以欧裔美国人还是认输的好，亚洲人就是更聪明。

事实很可能并不是这样。至少传统的智力测试显示的结果并不是这样。赫恩斯坦和莫瑞，拉什顿和詹森，菲利普·弗农和理查德·林恩以及其他人都曾在报告中指出，亚洲人具有智力优势。不过弗林已经证实，这些报告之所以得出这样的结果，是因为研究人员使用了过时的智力测试版本，而不是最新的版本，并且研究样本的规模较小，也不具有代表性。所以，使用过时的智力测试版本就会错误地提高亚洲人的智商。弗林回顾了16项不同的研究，这些研究的结果都基本一致，大多数研究结果表明，东亚人的智商略低于美国人的智商。

毋庸置疑的是，亚裔美国人取得的成就远远高于他们测得的智商显示的成就水平。亚洲人在成绩上的成就，更多的是靠辛勤的汗水换来的，而与智力无关。

亚洲人的智力成就来源于刻苦努力、积极和坚持不懈

哈罗德·史蒂文森及其同事对三个不同城市的学生进行了研究，这三个城市的社会经济水平非常接近：日本仙台、中国台北，以及美国的明尼阿波利斯市。研究人员从一年级和五年级随机挑选出一部分学生作为样本，然后对他们进行了智力测试以及阅读和数学水平测试。我们并不清楚他们对于三个不同人群的智力测试是否具有可比性（不过研究人员认为有可比性——并且给出的论据也比较有说服力）。尽管如此，一年级时，美国学生在绝大多数智力测试中的分数都高于日本学生和中国学生。作者认为，这是因为美国的家长在子女进入学校以前付出了更多的努力，激发子女的智力发展。不论是什么原因导致了美国一年级学生的表现更加优秀，到了五

年级时，美国孩子在智力的优越地位几乎丧失殆尽。从这些事实中我们了解到，不管一年级时谁比谁更聪明，五年级时美国孩子在智力上的优势已经拱手让给了亚洲孩子。

不过这项研究中真正值得注意的研究结果是，亚洲学生的数学成绩远远高于美国学生。使用同样的试题对日本、中国台湾地区和美国的孩子进行测试，到了五年级，中国台湾孩子的数学成绩比美国孩子高出将近1个标准差，而日本孩子的数学成绩比美国孩子高出1.3个标准差。更令人吃惊的是，在一项范围更大的研究中，史蒂文森及其同事研究了中国大陆和中国台湾地区以及日本很多所不同学校中五年级学生的数学水平。结果显示，这些亚洲孩子的数学成绩没有太大的差别，这些学校的数学成绩大致在同一水平上，而美国学校之间的差距就比较大。但是美国数学成绩最好的学校，其水平仅相当于亚洲最差的学校！

智商不是差距存在的关键所在，使美国学校或美国儿童与亚洲学校或者亚洲儿童有所区别的是学习的积极性。

让我们先从学校说起。日本孩子每年上学的时间是240天，而美国孩子上学的时间大约是180天。亚洲的学校可能更好，但美国学校中亚裔学生的表现说明，问题的关键在于学生学习的积极性。

于1966年发表的关于美国教育机会平等性的科尔曼报告，测量了随机选取的大量美国孩子的智商，弗林跟踪记录了这些孩子直到平均年龄36岁时的情况。其中，有东亚血统的美国人在智力测试非言语部分的得分是100分，语言部分的得分是97分，所以他们的总体智商略低于有着欧洲血统的美国人。

尽管在智力测试上略逊一筹，1966年这个班级中的华裔美国人在基础教育中留级的比例比其他孩子低50%。不仅如此，这些华裔儿童小学时成就测试的成绩也略高于欧裔小学生。到了高中，华裔美国人在成就测试中的成绩比欧裔美国人高出1/3个标准差。在同一智力水平上，华裔美国人在典型的成就测试中的成绩比欧裔美国人高出1/2个标准差。这种优势在数学测试上尤为明显。在微积分和分析几何测试中，华裔美国人的成绩比欧裔美国人高出整整1个标准差。当进入高中三年级时，华裔美国学生的SAT考试成绩比具有相同智商的美国学生高出1/3个标准差。

到32岁时，那个班级中的华裔美国人的付出得到了双倍的回报。为了获得成为专业人士或管理或技术工作所必需的教育资格证书，华裔美国人的智商至少要达到93分，白人则要达到100分。更重要的是，在智商达到了这些标准的人中，有78%的华裔坚持不懈最终拿到了资格证书，并且得到了期望的工作，而白人的这一比例则为60%。最终的结果是：55%的华裔拥

有较高社会经济地位的工作，而只有1/3的白人获得了这样的工作。日裔美国人的数据大概介于华裔和白人之间。

弗林在对东亚人的大量研究中也发现了这一现象，即亚裔美国人在成就测验和职业方面取得的成就高于根据他们的智力预测出的结果。

请注意，亚裔美国人的超预期成就，也就是他们智力测试的结果与其学习成绩之间的差异，足以证明从学前班到十二年级进行的成就测验以及SAT考试绝不是智力测试的代名词。这些考试反映出的智力成就不同于智力测试测量的记忆力、感知力和推理能力。另外，亚裔美国人的超预期成就也证明了，学业成就比智力更能准确地预测人们最终取得的社会经济成就。

最近，弗林研究了1966年那个班级中的亚裔学生的子女。既然我们知道成长于较高社会经济地位的家庭与高智商有关，我们就可以认为这些孩子的智商不仅应高于他们父母的智商，而且应高于全部人口的整体水平。实际情况的确如此。这些亚裔孩子上学之前的智商就比白人孩子的平均智商高9分。但是后来他们中的大多数都进入了普通的美国中学读书，我们可以想象，这对于他们的智力发展并不理想。事实上，他们的平均智商持续下降，直到成年时，他们的智商仅比白人的平均智商高3分。

请注意，我将亚洲人的成就称为“超预期成就”可能有些不合适。我有一位韩国朋友在美国生活刚满一年，他的孩子在公立学校里念书。有一次，我在他面前使用了“超预期成就”这个词。“你所说的‘亚洲人的超预期成就’是指什么呢？”他争辩道，“你应该说是‘美国人的成绩不良’！”他告诉我，有一次他参加女儿学校的期末典礼，着实吃惊不小。他发现，有一个奖项竟然是颁发给完成全部家庭作业的学生的。他的女儿就是两名获奖者之一。在他看来，对完成家庭作业的学生进行奖励简直和对吃午饭的学生进行奖励一样可笑。亚洲人显然认为完成家庭作业是理所应当的事。他坚持认为这种现象代表的是美国人的成绩不良，这是合理的。在他眼里，取得好成绩是一个人的本分，而绝大多数美国学生对待学习的态度却多少有些懈怠，这也是有道理的。

这位韩国朋友的困惑触及了理解亚洲人成就“魔法”的关键。

亚洲人和亚裔美国人的成就并没有什么神秘的，完全是靠刻苦努力得到的。20世纪80年代，日本高中生平均每天学习3.5小时，现在可能时间更长。东南亚“船民”的子女读高中时每天平均学习3小时，美国高中生整体上平均每天学习1.5小时。（底特律的研究显示，黑人八年级学生平均每周学习2小时。当然，不能完成作业至少有一部分原因是学校对学生的要求不高。）

亚洲人和亚裔儿童学习如此刻苦的原因也没有什么神秘可言。孔子在2 500多年前就已经把这件事说清楚了。他把能力的来源分为两类：一类来自自然——是上天给予的；另一类则是刻苦努力的结果。

如今亚洲人仍然相信智力成就——至少是优异的数学成绩——主要是通过刻苦努力而取得的，而欧裔美国人则更愿意相信智力主要是天生的能力或者得益于好老师。亚裔美国人在这件事上的态度介于东亚人和欧裔美国人之间。

和西方人（包括欧裔美国人）相比，亚洲人和亚裔美国人还有另外一个积极性上的优势。如果他们在某一件事情上表现欠佳，他们就会更加努力地钻研。加拿大的几位心理学家将一些日本大学生和加拿大大学生带到一间实验室，让他们参加一项关于创造力的测试。测试进行了一段时间以后，研究人员向参与者们致谢，并告诉他们测试的结果。不考虑他们实际上究竟完成得如何，研究人员告诉一部分参与者，说他们的表现非常出色，然后告诉另外一部分参与者，说他们的表现不令人满意。然后研究人员又向参与者们分发了一个类似的创造力测试，并告知他们愿意花多长时间来完成这个测试都可以。在第一个测试中被告知表现优秀的加拿大大学生会花更多的时间完成第二个测试，而在第二个测试上花费了更多时间的日本大学生，都是在第一次实验中被告知表现欠佳的学生。

面对失败坚持不懈是亚洲人自我提升传统中的重要元素。而在自我提升的过程中，亚洲人习惯于面对批评，而西方人的回应方式往往是逃避或者憎恶。日本的教师在任教10年以后还会有人在教室里旁听观察，其他教师会给出有关教学方法的各种反馈意见。在日本，人们一般都认为没有很多年的经验根本不可能成为一名好教师。而在美国，我们往往采用投硬币的方式来决定哪位老师管理哪个班级，我们假设他们一开始就能够胜任教师的工作，如果他们不能胜任，那是因为他们还没有掌握要领。

亚洲人能够最大限度地开发他们天生的智力还有一个更为重要的原因，那就是他们的文化——由他们的家庭传承给他们——要求他们这样做。以中国文化为例，早在2 000多年前中国人就非常重视学业成就。一个男孩如果聪慧好学，并且在官府主持的考试中取得优异的成绩，就极有可能为自己谋得一个收入丰厚的显赫官位。这不仅给他的家族带来了荣耀和财富，连他的同乡也能跟着沾光——是家人和父老乡亲的希望和期待在激励着他用功读书。和西方相比，早在2 000多年前，中国人就开始将教育作为升入更高社会阶层的云梯。

所以，亚洲家庭能够更成功地促进子女取得优异的学习成绩，因为亚洲家庭比美国家庭更能对子女施加强大的影响力——他们更加看重学业成就。

东方人的相互依靠和西方人的各自独立

为什么亚洲家庭能够施加强大的影响力呢？在这里我需要暂且退回一两步，先说说亚洲社会和西方社会在某些方面的巨大差异。亚洲人彼此之间互相扶持，更具集体主义精神，而西方人则比较独立，更倾向于个人主义。这种东西方之间的差异(注)早在至少2 500年前孔子和古希腊的时代就出现了。

孔子强调，严格遵守各种角色之间的关系是社会存在的基础，这些关系包括君臣关系、夫妻关系、父子关系、兄弟关系和朋友之间的关系。作为所有东亚社会的原型，中国社会是一个农业社会。在这样的社会中，尤其是严重依赖灌溉的农业社会，农民彼此之间必须要和睦相处，因为经济活动要求他们必须进行紧密合作。这样的社会往往等级森严，权力从上至下传递。个人的行为受社会束缚与限制的程度很深。家族，是中国社会尤其关键的构成单元。即使是现在，服从长辈的意志在很大程度上仍然是将家庭成员维系在一起的重要纽带。

家庭的传统角色仍然是第二代甚至第三代亚裔美国人与父母的关系之间一个不容忽视的因素。我的一些亚裔学生曾经告诉我，他们很想从事心理学或哲学方面的研究，但那是不可能的，因为父母想要他们成为医生或工程师。而对我的一些欧裔学生来说，父母对于他们未来职业的期望就好像父母的艺术品位一样，与他们毫无关系。

古希腊人的传统促生了一种完全不同的社会关系模式。希腊的经济基础不是建立在大规模的农业活动上的，而是贸易、狩猎、钓鱼、畜牧、海盗掠夺以及小规模的与农业相关的生意，例如葡萄栽培或生产橄榄油。所以这些活动都不要求人与人之间形成固定模式的紧密联系。所以，希腊人大都独立，可以享受不受社会限制和束缚的自由。他们有很大的自由空间展现他们的才华，满足自身需求。个性得到了张扬，并成为评论和研究的主题。罗马社会延续了古希腊尊重独立及个人主义的传统。之后一段漫长的岁月里，欧洲的农民很可能与中国的农民一样不具备个人主义的精神。不过后来的文艺复兴和工业革命，又使西方文化中的个人主义得以延续，甚至加速了个人主义的发展。

只接受过欧洲文化熏陶的人很难理解，在东方个人的成就关乎整个家族的利益，而不仅仅是个人的骄傲和地位。就像中国古代等待走入仕途的生员一样，一个人的成功关系到整个家族的利益（既包括经济利益，也包括社会利益）。尽管成就能够给个人带来成就感，但绝不仅仅是一件给自己带来利益和荣誉的事情。

不仅如此，对于家庭来说，为家庭取得成就似乎比为自己取得成就更能激励人追求成功。如果我，一个按照自己意愿行事的西方人，为了给自己带来荣耀或财富而选择追求成功，这是我个人的决定。如果我认为我的才能不够或我不想努力用功，我就可以选择退出忙忙碌碌的生活或工作。但是如果我和我的家庭被紧密地联系在一起，并且在自我谋生的同时还要满足家庭对成就的需求的话，我就别无选择了，只能在学校和未来的职业生涯中拼尽全力。而这种需求也是合理的，因为我已经完全清楚只要努力就可以取得成就，而不是依靠天生的才能。

相对于欧裔美国人，亚裔美国人的成就优势很有可能会继续扩大。在1968年以前，来到美国的亚洲移民很可能天生并不比他们在国内的同胞聪明。不过20世纪60年代的《移民法》使得专业人士移民到美国变得相对容易一些，而非专业人士则相对困难一些。因此，亚裔美国人来到美国之初，就比欧裔美国人拥有文化上的优势，他们不光是专业人士或身居管理职位，同时他们还是东亚人。他们所属的社会阶层和所接受的文化都有利于子女在教育和职业方面取得最大的成就。这也会使他们的子女具备遗传优势。（不过这种基因优势很可能并不大。在下一章我们就会看到，对于受瓶颈制约的人的后代来说，环境瓶颈并不会对他们的智力产生多大影响。）

全局型思维与分析型思维

东西方的文化差异不仅仅使两者在智力成就上有着量的差距，在思维方式上也产生了质的差别。东亚社会的有效运转都依赖于将个人的需要和行动与他人结合在一起。和谐是中国2 500年来所有社会关系的关键词。而西方社会的有效运转则不依靠与他人打交道，而是个人采取独立的行动，不考虑他人的意愿。

这些社会差异导致东方人在思维方式上形成了一定的特点，我将其称为“全局型思维”。东方人往往关注的是一个很广的范围内的事物；他们更关心事物彼此之间的关系和相似之处；他们使用辩证思维去推理，这其中包括在完全对立的观点中找出“折中的道路”。而西方人的感知和思维方式则是“分析型”的，也就是说西方人关注的是整个环境中的局部，也就是他们能够对其施加某些影响的物或人。他们关注这些局部往往是为了进行分类并根据其行为构建模型。他们通常使用逻辑定律进行推理。

顾及他人也就意味着东方人的感知都向外集中于广泛的社会环境，并因此集中于物理环境。增田贵彦和我向一些人播放了一些展示水下场景的动画短片，然后让参与者说出他们都看到了什么。我们来看一下图8–1，这是从这些片断里截取出来的一幅定格的画面。美国人主要将注意力放在明显的物体上——例如，迅速游动的大鱼。通常他们的第一反应都是：“我看

见三条大鱼游向左边；它们的白肚皮上有粉色的斑点。”

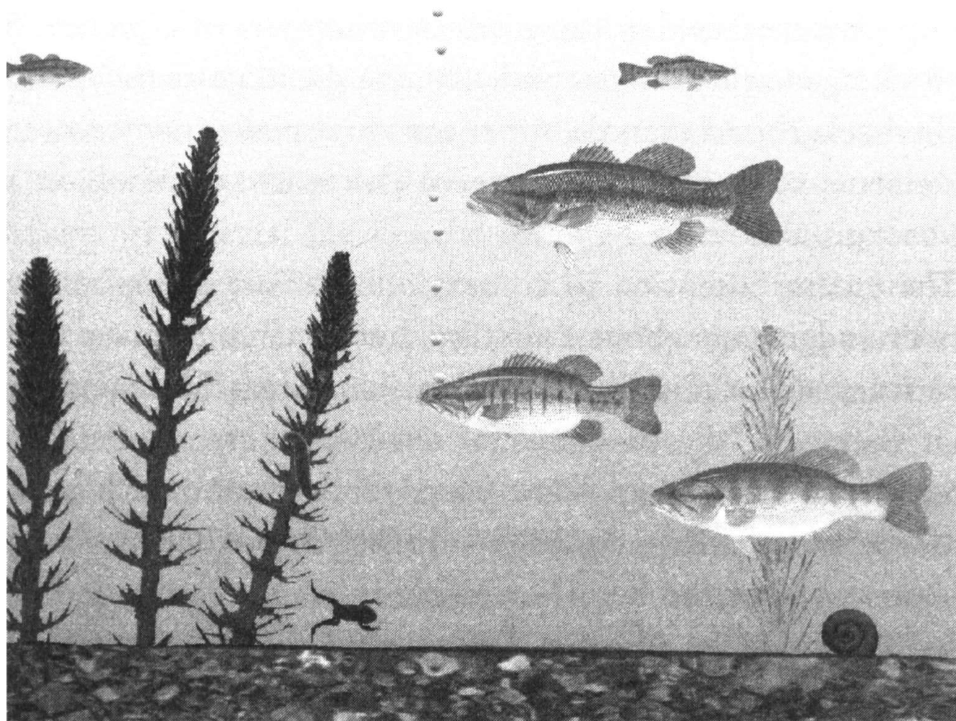


图**8–1** 从一段彩色动画中截取的画面

日本人看到了这个环境中更多的东西——岩石，水草，静止的生物（例如蜗牛）。典型的第一反应是：“我看到的好像是一条溪流，溪水是绿色的，我看见底部有岩石和贝壳。”除了观察大环境以外，日本人还注意到了环境中的具体物体与整体环境之间的关系。例如，他们往往会注意到一个物体与另一个物体挨得比较近，或者一只青蛙正在跃上一株植物。日本人说出的环境中的细节比美国人多60%。

在另一项研究中，增田将一幅卡通图片展示给不同的人看，这幅图片的中心有一个人，这个人周围簇拥着其他一些人。增田要求观看者判断这个位于正中的人的情绪。与美国人相比，日本人往往会更多地根据周围人的表情来判断这个人的情绪。

亚洲人和西方人看到了不同的东西，是因为他们关注的东西不一样。我和我的同事让观看者佩戴了一些仪器，这些仪器能够观察到观看者在千分之一秒内所关注的图片的部分。中国人比美国人花了更多的时间观察整个背景，并且眼睛在明显的物体和背景之间来回转动的次数也多于美国人。

更注重整体使得东亚人能够在一定的情景中对因果关系做出正确的判断，而美国人则往往做出错误的判断。社会心理学家已经发现了他们称之为“基本归因错误”的现象。人容易忽略某种行为发生的重要社会和情境原因，而将这一行为认定为行动者的特征——个性特征、能力或态度。例如，某个科目的任课老师或某项心理研究中的实验人员让某位作者写一篇支持死刑的文章。当读到这篇文章时，美国人往往认为作者表达的观点一定是他自己的看法。即使是实验人员刚刚要求他们按照指定的观点写一篇支持的文章，他们仍然认为所读文章的观点就是作者自己的观点。在这种情况下，韩国人往往会做出正确的判断，他们不会认为文章的作者实际上也持有他在文章中所表达的观点。

东亚人这种把握全局的特点早在中国古代就开始形成了，这使得他们能够在更大的背景中理解某个事物，所以他们就有可能发现磁力和声学的原理，以及潮汐产生的真正原因（这个连伽利略都没有发现）。相反，亚里士多德的物理学则完全关注物体本身的特征。在他的理论体系中，石块落入水中后就会沉底是因为自身存在重力，一块木头浮在水上是因为自身存在浮力。然而物体自身当然不具备浮力和重力的特征，只有在物体和其他物体的联系中才存在浮力和重力的概念。

虽然中国古代的物理知识更正确，中国古代在科技上的成就也远远超过古希腊，但创立了形式科学的还是希腊人，这主要有两个原因。

第一，希腊人将注意力都集中在物体上，他们关注的是物体的特点和归类。为了理解物体的运行状况，他们就推测那些控制物体运行的规律。规律和分类是科学的基本组成部分。没有规律和分类，就不存在具有概括性的显式模型来验证世界。不论事物多么复杂，都一定存在规律。

第二，希腊人发明了形式逻辑。据说，亚里士多德听腻了市场上和政治集会中那些漏洞百出的论点，于是发明了形式逻辑，以避免论点存在逻辑上的漏洞。不管怎么说，逻辑确实在西方实现了这样的功效。

中国人好像从来都对逻辑不太感兴趣，事实上，逻辑只在公元前3世纪短暂地出现过一次，从来没有形成过一个学科。希腊人能够发明形式逻辑完全是因为有被社会普遍接受的辩论习惯。在中国古代，以及如今在东亚的大部分地区，提出不同意见是有风险的——你或许会因为提出了不同意见而在不经意间树敌。所以东方的抽象推理模式并没有朝逻辑的方向发展，而是朝着辩证的方向延伸，注意从对立的观点里求得“折中的道路”，并且强调不同观点的相互融合。

就像定律、分类和显式模型一样，形式逻辑对于科学来说是一个极其有用的工具。不过希腊人对逻辑辩论的热爱走向了极端。他们不承认零的概

念，因为按照他们的推理，零就等于“不存在”，而“不存在”怎么可能存在呢？还有芝诺[注]著名的悖论，就是逻辑不受控制的结果。（例如，运动是不可能存在的。如果一支箭要到达靶子，就先要经过弓与靶之间一半的距离，然后再走剩下的距离中一半的距离，不断进行以至于无穷，所以箭永远也射不到靶子上。在我们现代人看来这种逻辑多么可笑，但当时的希腊人却认为这是个无法解决的问题。）

社会风俗和思维习惯往往根深蒂固，所以如今东方和西方在社会和认知方面的差异仍和古代没有太大的差别。西方人更强调规律、分类和逻辑，而东方人往往更关注事物之间的联系和辩证推理。我和我的同事在实验中得到的结果也印证了这一点。

我们给参与者提供三个单词，例如牛、鸡和草，然后让他们将其中两个词归为一组。我们从东方人和西方人那里得到了非常不同的答案。美国人更多地将牛和鸡归为一组，因为它们都是动物，也就是说，他们同属于一个分类学范畴；而亚洲人更关注关系，所以更多的将牛和草归为一组，因为牛吃草。

我们还向美国人和亚洲人展示了一些三段论，并要求他们判断这些三段论的结论是否成立。我们发现，当判断一些用抽象概念表述的三段论——所有的A都是X，一部分B都是Y——的结论是否成立时，亚洲人的判断和美国人一样正确；而当参与者熟悉三段论中的内容时，亚洲人往往容易被误导。亚洲人往往认为如果三段论的结论本身不合理，那么即使这个结论是根据前提推导出来的，也是不成立的。（例如，所有的哺乳动物都冬眠/兔子不冬眠/所以兔子不属于哺乳动物。）他们容易将不是根据前提推导出来的但却合理的结论，判断为正确的结论。

最后，我们还发现，美国人有时在推理上也会犯错，就是因为他们也采取了古希腊人典型的“超逻辑”立场。我和我的同事发现，比起不存在与之相悖的命题，如果有一个给定的合理的命题与一个不怎么合理的命题相悖，美国人更倾向于认为与一个不怎么合理的命题相悖的命题是真的。美国人认为，如果两个命题之间明显相悖，更合理的那个命题肯定是真的，而不怎么合理的那个命题肯定是假的。而亚洲人犯的错误正好相反，他们认为，明显相悖的两个命题中，更不合理的那个肯定是真的——因为他们乐于在两个对立的命题之中都找出一些合理的内容。

这些感知和认知差异所反映的，是东方人和西方人大脑活动的差异。例如，当中国人看水下场景的动画片时，大脑中与背景和全局相对应的区域，活动得比美国人更频繁。相反，大脑中对应明显物体的区域的活动不如美国人频繁。另一项有关大脑功能的研究显示，美国人在忽略掉问题的

周遭环境时更容易做出判断，而东亚人在全局环境中考虑问题时更容易做出判断。大脑的活动也证明了这一点：当人做出非优先的、更加困难的决定时——也就是需要美国人把问题的前后关联考虑进来做出判断，以及东亚人忽视问题的前后关联时做出判断，与注意力控制有关的额叶和顶叶皮质区域的活动更加频繁。

那我们如何知道这些感知和思维方面的差异是由社会造成的，而不是由基因造成的呢？主要的原因有两个。第一，在我们进行的几项研究中，我们比较了亚洲人、亚裔美国人和欧裔美国人。在所有的研究中，亚裔美国人的感知和推理方式刚好介于亚洲人和欧裔美国人中间，但是通常更接近欧裔美国人。第二，众所周知，中国香港是一个东西方文化融合的地区，中国习俗与西方习俗交织在一起。我们发现香港人推理的方式介于中国人和欧裔美国人之间。当我们要求香港人对鱼的行为进行因果归因时，如果他们看到的是庙宇和龙的图片，其推理方式就接近中国人，而当他们看到的是米老鼠或美国国会大厦的图片时，其推理方式就更接近西方人！

东方工程师和西方科学家

从东方人和西方人在社会倾向和思维模式上的差异，就可以推断出他们适合做工程师，还是更合适做科学家。

每个人都听说过这样的陈词滥调：日本人只能成为优秀的工程师却做不了优秀的科学家。这可不完全是对日本人的成见。日本人在工程学方面的卓越才能是美国工业中的一个奇迹。我的一些教授工程学的同事和聘用工程师的朋友都告诉我，亚裔美国人中工程师的比例很高，而且他们的平均水平大体都优于欧裔美国工程师。

尽管如此，在20世纪的最后10年里，有44项诺贝尔科学奖的获得者都居住在美国，其中绝大多数都是美国人，只有一位是日本人。这和科研资金的多少并没有太大关系。在过去的25年里，日本人进行基础研究的费用只是美国的37%。虽然德国进行基础研究的费用只有日本的一半，但德国在20世纪90年代获得了5项诺贝尔奖。中国和韩国直到最近才渐渐富裕起来，所以现在就说他们的国民在基础研究中能够取得何种成功还为时尚早。不过我们现在可以发现一些通往科学成果的路上存在的障碍，而这些障碍可能会挡住所有具备全局型思维、倾向于彼此依靠的民族。

第一，东西方的一些社会差异对西方取得科学进步更为有利。与西方国家相比，日本在很多方面的社会等级划分更加严格，也更尊重前辈，所以更多的研究经费会划拨给那些已经上了年纪的、不再有什么成果的科学家。我相信，西方对个人成就的重视和对个人抱负的尊重更有利于人取得科学成就。在实验室里长时间工作并不会给科学家的家人带来什么益处，却能

给科学家本人带来声望或荣誉。在西方，辩论是司空见惯的事，也被认为对于科学事业来说是非常必要的，而在东方，这样做则是非常无礼的行为。一位日本科学家最近说：“我在华盛顿的卡内基学会工作，我认识两位卓越的科学家，他们两个人私下里是好朋友。可一旦涉及工作，他们就会进行非常激烈的辩论，有时甚至在报刊上公开向对方开火。这种事在美国可能会发生，但在日本绝对不会。”

第二，儒家传统并不赞同“知识本身是宝贵的”这种观点，而日本和韩国在这方面也受到了儒家传统的影响，这和古希腊的哲学传统形成了鲜明的对比。古希腊的哲学传统认为这样的知识在所有的知识里是最宝贵的。（我要强调上一句话中出现的哲学传统几个字。在《理想国》里有一个有趣的故事，讲的是一个雅典商人痛斥苏格拉底追求没有实际价值的知识。他说，年轻人追求这样的知识很令人赞赏，而上了年纪还追求抽象的知识就很令人反感。）

第三，西方人比东方人更习惯于在现实生活中运用辩论的智力工具，即逻辑。即使西方人有偶尔过度使用逻辑的习惯，这也对科学非常有帮助，不管它在现实生活中是多么不好用甚至可笑。在正式的科学、法律和政策分析的论述中，西方人最常使用的修辞就是逻辑，其中包括概述讨论的主题，与主题有关的种种问题，具体的假设，对假设进行论证的方法，相关事实的讨论，对有可能出现的相反论点进行驳斥，以及对结论的总结。这种对辩论模式的训练从幼儿园就开始了：“我喜欢这只泰迪熊，因为……”西方对这一修辞手段的普遍使用也许与其辩论和形式逻辑的渊源有关，但这在东方并不常见。我从我的东亚学生身上发现，标准修辞形式是他们在攻读博士学位时最不愿学的科目。

最后，好奇心也起到了一定的作用。不管是出于什么原因，西方人似乎比东方人更富有好奇心。西方人探索地球并全身心地投入到科学之中，西方人将哲学研究视为人类的本质。我不知道为什么会这样，不过我可以猜测其中的一个原因。我们都知道西方人一直不断地构建世界的因果模型。实际上，美国老师经常认为，因为经商而居住在美国的日本人的子女缺乏分析能力，因为他们不会搭建这种因果模型。搭建显式模型可能会出现的结果就是得出意外的结论。得出的结论可能与实际情况正好相反，这就促使一个人寻求更加正确的观点——也就更加激发了好奇心。

东方人典型的思维习惯并不会在他们通往科学成就的路上设置难以逾越的障碍。科学实践促生了我的认知——认为西方具有优势的思考模式，并且接受的科学文化越多，头脑就会越来越自然地养成科学思考的习惯。这样东方人就可以将他们与众不同的思维习惯沿着有利于科学探寻的方向加以改进。物理学里量子理论的矛盾之处令西方人极为头疼，但和东方人却很契合。尼尔斯·玻尔就将自己做出的量子假设归功于自己对东方哲学的深刻

理解。

目前，在与东方的友好竞争中，西方在科学上的优势或许正是他们的秘密王牌。不过，不要指望这样的优势能够保持多久。不要忘了，直到20世纪过了一多半的时候，欧洲科学家还在疑惑为什么美国就没有什么重要的科学发现呢。

---

1. “船民”，指20世纪70年代末期，被越南当局强迫乘船离境的难民。——译者注

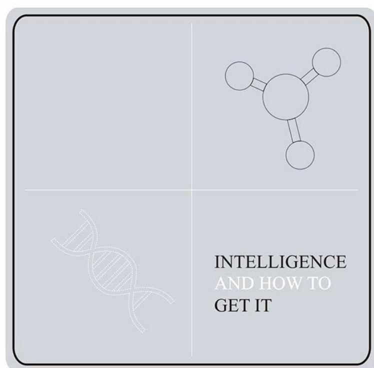
2. 在这一章的所有讨论中，全部的概括和总结都只适用于东亚人，而不是南亚人。我相信有些概括也适用于南亚人，但是证据——有很多证据——主要与东亚人有关。后面的内容中我谈到的“亚洲人”或者“东方人”都是指东亚人，例如中国大陆和港台地区的人、日本人、韩国人、新加坡人。儒家学说构成了所有这些国家的文化根基，并深深地影响了这些地方的文化。

3. 我所知道的所有有关亚洲人，或者亚裔美国人，与欧裔美国人之间的比较都显示，和种种测量言语智商的测试相比，亚洲人或者亚裔美国人在智力测试，尤其是例如积木图案这样有关视觉空间的测试中的得分更高，不过这两类测试的得分差距很小。赫恩斯坦和莫瑞认为，亚洲人的操作能力略强于欧裔美国人主要是基因所致，我们当然不能排除这种情况。不过另一方面，正如你们在这一章后面的部分将要看到的那样，亚洲人或者亚裔美国人的文化，更强调关注更广阔的视觉空间世界。另外，我的东亚同事曾经告诉我，东亚学校中教授的类似于智力测试中的空间问题要比西方学校里教的难得多。

4. 我所谈到的东方人和西方人的社会差异，以及由此造成的两者在认知方面的差异，都是对我在《思维版图》中提出的观点的概述。这本书介绍了欧洲人及欧裔美国人，与东亚人及东亚裔美国人在历史和现代不同时期、不同社会中的认知发展趋势。我推荐大家参考一下这本书，因为里面不仅包括了这一章中出现的绝大多数有关亚洲人和西方人的历史差异的论断，还详细介绍了大量有关现代人感知和认知习惯的研究。

5. 芝诺（Zeno），公元前5世纪前后古希腊埃利亚学派哲学家。

第9章
智慧的象征——犹太人



一个男人应该变卖所有财产去娶一位学者的女儿，也应该把自己的女儿嫁给一位学者。

——《塔木德经》

“犹太人”显然是这个世界上少有的智力贵族。

——马克·吐温，**1879**年的信件中

现在世界上拳击运动发展最好的国家是美国，但仔细研究美国所有4 000多名拳击手后就会发现，其中的优秀选手都是犹太人。

——乔·汉弗莱，美国拳击赛事播音员（**1930**年）

公元64年，犹太主教耶霍舒亚·本·盖姆拉颁布了一项法令，明确规定所有的男性都必须能够诵读《塔木德经》。在这项法令颁布了100年之后，所有犹太男性终于达到了这一要求。下一个在男性间普及了读写能力的民族大约出现在1 700年以后。

犹太人的杰出成就

最先普及读写技能的民族现在也在智力方面取得了极其辉煌的成就，这绝不是某种巧合。德系犹太人获得诺贝尔奖的比例远远高出犹太人在全世界人口中所占的比例，过度代表（overrepresent）的系数从50：1（诺贝尔和平奖）到200：1（诺贝尔经济学奖）不等。不过比较犹太获奖者在西方国家人口中所占的比例，甚至是在美国人口中所占的比例，可能更公平一些。在所有颁发给美国人的诺贝尔科学奖中，美国犹太人占到了27%~40%（取决于我们是将犹太人定义为父母双方都是犹太人的人，还是将只有一半犹太血统的获奖者也定义为犹太人）。犹太人在美国总人口中的比例不足2%，所以过度代表的系数就是15：1（使用更保守的犹太人的定义）。在图灵奖的美国获奖者中也出现了这种过度代表的现象。菲尔茨奖的美国获奖者中有26%~34%都是犹太人。

在美国，常春藤联盟学校中有33%的学生都是犹太学生；还有差不多同样比例的犹太人在名牌大学里任教；最高法院中有30%的书记员是犹太人。过度代表的系数都达到了15：1甚至更高。

犹太人的成就还不只局限在智力方面。根据波兰1931年进行的人口普查，犹太人占总人口的9.8%，但是掌握着这个国家22.4%的财富。第一次世界大战以后的4年里，有超过70%的营业执照都是颁发给犹太人的。到1929年时，45%的大中型商业企业都掌握在犹太人的手里。到1938年，这一比例增加到了55%。20世纪30年代中期到末期，纺织、化学品、食品、交通运输、建筑材料和造纸行业的所有者中，犹太人占据了一大半。

所以，犹太人在很多以高智力决定优势的领域都取得了极其辉煌的成功，包括商业和贸易。

犹太人的平均智商是最高的

犹太人给人的印象通常都是智商高，当然这是有统计数字可依的。智力测试显示的也是同样的结果。在所有有可靠测试数据的民族中，犹太人的平均智商是最高的。根据大多数报告，德系犹太人的平均智商比白人的平均智商高出2/3~1个标准差。也就是说，他们的智商是110~115分。

犹太人在智力领域过度代表的系数，实际上超出了以犹太人的智力作为参考得出的过度代表系数。让我们暂且将取得最高智力成就——诺贝尔奖或类似奖项——的平均智商定为140分或140分以上。如果我们假设犹太人的平均智商是110分[注]，那么犹太人在智商为140分的人中的过度代表系数只有6：1。这远远低于我们发现的至少15：1的过度代表系数，也远远低于我们按照父母其中一方为犹太人的宽泛定义计算出的过度代表系数。即使我们假设犹太人的平均智商是115分，他们实际上取得的成就也高于

按照智商预测的成就水平。我们也可以暂且将130分定为就读于常春藤联盟的学生、名牌大学里的教师和最高法院书记员的平均智商。假设犹太人的智商是110分，那么犹太人在智商为130分的人中的过度代表系数只有4：1，这远远低于他们实际获得的大约15：1的过度代表系数。假设犹太人的智商是115分，那么犹太人在智商为130分的人中的过度代表系数只有7：1，这也远远低于实际的系数。如果我们认真看待这些数字——实际上我建议大家不要持有太多的保留意见——那么犹太人就不只是取得了很高的成就，而且他们取得的成就大于根据他们的智商所预测的成就。

值得注意的是注，西班牙系犹太人（主要是北非人）的平均智商显然并不比非犹太人的智商更高，而且比德系犹太人的平均智商低得多。即使是以色列的西班牙系犹太人和东方犹太人，情况也是如此。

犹太人的高智商源自遗传？

德系犹太人与西方人在智商和智力成就方面的差距是基因造成的吗？对于犹太人具备智力上的优势，当然不乏与遗传基因有关的解释。这里我就介绍5个最常被援引的解释。

1．迫害的馈赠。一个非常古老的遗传论解释就是，欧洲犹太人中最笨的人受到的迫害最大，因为根据推测，这些不聪明的人不知道如何从敌人手下逃走。从基因库里超出正常比例地淘汰掉最笨的犹太人，就会使剩下的犹太人的平均智商有所提高。这种解释有两个问题。第一，我们不能完全确定最笨的犹太人被集体杀戮的可能性更大。我们还有理由认为，更有可能遭到打击的应该是最富有最聪明的犹太人，因为他们更容易引起别人的注意。第二，我们还不清楚从基因库里超出正常比例地淘汰掉笨的人，是否就会取得如此大的效果。这一解释认定的现象被称为遗传“瓶颈”——在不同寻常的环境状况下，从繁殖种群中除去某些基因型所导致的后果。（瓶颈的经典例子就是，当某个地区的某个繁殖种群中的一小部分离开了这个地区时，只会从原来的基因差异中带走有限的一小部分。）尽管如此，即使我们假设智商具有很高的遗传度，瓶颈现象也不能合理地解释犹太人智商较高的原因。因为要将位于智商分布最底部的15%的人完全去除，随后的一代人的平均智商才会提高1分。只有经过很多事件将大量位于基因库下游的人完全从基因库中清理出去，才会使遗传基因有所变化。

2．尼布甲尼撒二世的偏爱。遗传学家西里尔·达灵顿曾经提出，巴比伦囚禁使得犹太人的智商有所提高。公元前586年，耶路撒冷被尼布甲尼撒二世攻占。根据《圣经》记载，尼布甲尼撒二世“将耶路撒冷的众民和众首领，所有勇士，连同一切木匠、铁匠都掳了去，除了国家中极贫穷的人以外，没有剩下的”（《列王记下》，第24章第10~14节）。这一假设再

次指出，被留下的相对不聪明的犹太人可能会到处漂泊，加入其他的宗教。所以当犹太人返回耶路撒冷时，他们就不会与被留下来的不聪明的犹太人有任何瓜葛了。除了认为不聪明的人离开了他们的宗教以外（这在很大程度上只是推断），这一理论还建立在一个单一的瓶颈假设的基础上。即使假设智力的遗传度非常高，并且进一步假设，智力与财富状况在古代是高度相关的，从繁殖种群中将最穷的人去掉一大部分也不会使下一代人比上一代人更聪明一些。

3．嫁给学者。另外一个非常流行的遗传理论就是，（极其聪明的）商人和生意人的女儿很有可能嫁给（极其聪明的）学者或拉比（犹太教负责执行教规、律法并主持宗教仪式的人）。这样的强强联合带来的不仅仅是财富的增加，他们的后代也更有可能生存繁衍下去。坚持这种理论的人有时会搬出《塔木德经》中有关嫁给学者的指令来。不过有人认为，不管是贫穷的犹太人，还是富裕的犹太人，都不愿意把女儿嫁给身无分文的学者，并且富人有可能更青睐生意人。另外，这种推测出来的对智力有利的结合，所涉及的人数肯定只是全部人口中非常小的一部分，所以我们就会怀疑，这样是否能够把人口的平均智商提高到这么高的水平。

4．不能阅读并理解《塔木德经》。政治科学家查尔斯·莫瑞曾经指出，仅具备读写能力恐怕还是不能够阅读复杂的《塔木德经》并理解它的意思，更别提对原义进行解析和诠释了。没有达到如此高水平的读写能力的人可能就逐渐脱离了群体，只剩下最聪明的人留下来繁衍后代。莫瑞的理论很有趣，但它仅仅是推测而已。

5. 职业压力。人类学家格雷戈里·科克伦、贾森·哈迪以及亨利·哈平丁曾经提出过迄今为止最为详尽的有关德系犹太人智商的理论。

德系犹太人第一次出现在欧洲的历史记录中，是在公元9世纪。从很早开始，他们就从事一些要求读写技能、数学能力和高智商的职业，包括金钱借贷（当时《反高利贷法》禁止基督徒放债），以及作包税人和地产管理（尤其是在东欧）。从事这些职业的人都发了财。财富意味着后代生存下来的概率更大。所以聪明人的后代比不聪明的人的后代多，整个犹太人口的平均智商就这样逐渐得到了提高。

科克伦及其同事还提出了一个德系犹太人智商提高的具体途径。首先，德系犹太人极易患上某些疾病，例如，泰—萨克斯病、尼曼–匹克病和葡糖脑苷贮积症。这些疾病与神经细胞里所谓的鞘脂类的含量有关系。鞘脂类构成了绝缘外鞘的一部分，绝缘外鞘能够使神经细胞发出电信号并促进神经细胞树突的生长。神经细胞里这种鞘脂类的含量一旦过高可能致命，至少有可能引发某些导致绝育的严重疾病。

但是为什么鞘脂类的含量过高就能使未患病的人口的智商有所提高呢？在这一点上，科克伦等人使用镰状细胞性贫血症做了一个类比。拥有两份镰状细胞基因的个体（一份来自父亲，一份来自母亲）就会患上镰状细胞性贫血症。但是只有一份镰状细胞基因的个体就会对疟疾形成免疫力。这对于西非人是很有好处的，因为他们在原来的居住地就饱受疟疾的侵袭，迄今为止，他们是世界上最有可能拥有镰状细胞基因的人。

科克伦将这种情况和德系犹太人的高智商进行类比。带有两份鞘脂类含量过高的基因的人，就会患上严重的疾病或者因此丧命，但是对于只有一份鞘脂类含量过高的基因的人，这只会造成神经细胞内鞘脂类的含量过高，但不会达到致命的程度。而鞘脂类的含量高能够促进神经信号的传播和神经细胞树突的生长。而据说神经元突起上的树突形成的分支越多越广，对学习和一般智力就越有利。

显然，鞘脂类确实能够促进神经传递和树突的生长。不过只有一个证据可以证明鞘脂类理论，那就是葡糖脑苷贮积症患者的智商比其他犹太人的智商还要高。在以色列，葡糖脑苷贮积症的患者大体都拥有地位极高的工作，其中不乏物理学家和技术人员。

和其他关于德系犹太人具有高智商遗传基因的推测相比，以镰状细胞的类比为模型的鞘脂类理论确实有一个优点。由此我们可以预测出，只有一份鞘脂类含量过高的基因的人，比没有这种基因的人的智商高。但是科克伦及其同事没有对这个预测进行证实，只是报告了葡糖脑苷贮积症患者极高的职业成就。我有点儿想不通，这一推测并不难证实，而且许多科学家在发表一个复杂的理论之前都会先证实其推测。

请注意，历史上犹太人很早就普及了读写技能，而科克伦的理论居然没有意识到这个事实与犹太人高智商之间的因果联系。读写技能使犹太人更易于从事一些职业，而这些职业又最终使得犹太人的智力得到了提高。科克伦的理论确实应该注意到，在职业优势还没有发挥它在遗传上的魔力之前，犹太人在智力方面并没有取得什么成就。在后来的数个世纪里，德系犹太人才不断取得越来越多的智力成就，并在19世纪中期开始向顶峰迈进。

还有一点对于科克伦的理论也很重要，西班牙裔犹太人并没有取得太多成就，这是因为他们并没有经过要求具备高智商的职业的过滤。现代的西班牙裔犹太人的智商实际上并不是特别高，但是信仰伊斯兰教的西班牙裔犹太人却取得了极高的成就。公元1150—1300年，15%的科学家都是犹太人——这与他们在全世界人口中的比例严重失调，甚至与他们在整个伊斯兰教世界人口中的比例失衡——而这些科学家绝大多数都是西班牙裔犹太人。科克伦等人对于西班牙裔犹太人达到的这一顶峰只给出了文化方面的

解释，而且这种解释和他们关于犹太人智力水平的遗传理论并不一致。

总之，虽然有很多关于犹太人高智商的遗传理论，但并没有多少人能够拿出令人信服的证据。

不同国家、民族和地区的智力差异

不同文化——即使是文盲率极低、经济极富活力的文化——取得的智力成就通常也有极大的差别，研究犹太人的高智商时我们需要考虑到这一因素。不过这些成就的高低与遗传基因无关。

公元1000年，世界上的知识分子主要是阿拉伯人和中国人。阿拉伯人讨论的是柏拉图和亚里士多德，中国的文人则精通各种艺术，而那时欧洲的贵族们正在潮湿寒冷的城堡里大快朵颐呢。智力分数牌为中国人和阿拉伯人（还有印度人）亮起，而欧洲人差不多得了零分。后来欧洲逐渐取得了辉煌的智力成就，一部分原因是他们乐于向更先进的文化学习，但这种智力中心发生转移的原因不可能是基因库发生了改变。

即使是在整个欧洲注，智力成就的宝座也数度易手。西班牙在摩尔人统治时期达到了智力成就的顶峰，但只是昙花一现，即使是在新大陆上发现了不计其数的金银的岁月里，也没有取得什么值得一提的成就。意大利北部在15世纪是所有艺术和科学的发电站，而那时英国的文化还是一潭死水。从1800年开始，英国几乎在所有领域都处于领导地位，在科学、哲学和文化方面的地位无人能够撼动。而意大利从1800年开始，只能算是跟随在罗马时代和文艺复兴时代身后的影子。苏格兰人在中世纪末期还是一群野蛮人，在战场上把自己涂成蓝色，靠刺杀的方式选出新的领袖。（莎士比亚把《麦克白》的故事背景放在苏格兰是有道理的。）但是到了18世纪，苏格兰人成了科学界和哲学界的领袖。而在20世纪之前，斯堪的纳维亚国家几乎没有取得什么重要的智力成就。

在美国，不同地区的智力成就也存在惊人的差异。东北部地区的人口从来没有超过南部地区，但是南部人在科学、哲学和艺术（除了音乐）方面取得的成就，根本不能与东北部人相提并论。整个20世纪，得克萨斯的人口是新英格兰人口的3～4倍，甚至得克萨斯的非拉美裔美国白人的人数，都超过了新英格兰的人口。但如果你翻翻名人录就会发现，得克萨斯人在科学和哲学方面几乎没有取得任何成就（不过最近的几十年里，得克萨斯人在文学、音乐以及艺术方面的成就开始引起大家的关注）。

在西方，和犹太人与非犹太人在智力成就上的巨大差距相比，国家与国家、民族与民族、地区与地区之间在智力成就上的差异更大。不过我们还

是需要弄清楚为什么犹太人与非犹太人在智力上存在差异。文化因素能够完全合理地解释这一差异吗？

文化成就了犹太人的高智商

和儒家传统一样，犹太传统同样重视教育的作用。正如奥地利小说家斯蒂芬·茨威格在第二次世界大战中写的那样："如果一个家庭来往的人中有一位教授，或学者，或音乐家，反正是那些在知识界享有一定地位的人，这就像给这一家子授予了荣誉称号，就好像这个人取得的成就使他们也变得尊贵起来一样。"

在另一个关键的方面，犹太人也拥有和儒家弟子极为相似的传统。在犹太家庭中，家庭成员之间的关系很亲密，家庭对个人提出的要求，个人很难拒绝。犹太家庭中的母亲，据说拥有的影响力丝毫不逊于任何中国父母，他们的形象甚至被编进了歌曲和故事里。而且她们主要是在教育和智力成就方面对孩子施加影响。

在非犹太人看来，犹太人的教育观有时候有些极端。心理学家西摩·萨拉森（Seymour Sarason）在一个贫困的社区里长大，家境非常贫寒。有一次他的一个表亲决定在高中踢一场足球赛，整个家庭都对这个决定表示了愤怒和震惊。他们担心一旦他受伤了，可能就无法去念大学了。这可以和《胜利之光》中的故事做个对比。萨拉森还回忆说，他的父亲没有钱，也没上过什么学，但是有一次他买回一本非常昂贵的百科全书式的字典，这给萨拉森留下了极其深刻的印象。他认识到，教育非常重要——即使是倾家荡产也要接受教育。

如今，在布鲁克林，与有些孩子交换足球卡不同，哈西德派的男孩交换的是拉比卡。

我们可以列出无数这样的故事来说明犹太人对教育、精神生活和成就的重视。当然他们看重的成就并不局限于智力领域的成就。过去犹太人除了在商业领域取得了成功以外，还非常重视体育方面的成就。20世纪初期，犹太人在拳击、摔跤和篮球（一名敌视闪米特人的解说员曾称，篮球是犹太人称王称霸的项目，因为运动员要能够狡猾地把球从别人手中偷走）等体育项目上取得的成就无人能及。如此看来，犹太人确实很注重成就。

这些证据虽然都显示了犹太人重视智力成就的文化观，但它们也只是故事而已。我们无法测量文化对于智力成就的影响，甚至也无法像雪莉·布莱斯·希斯那样，系统地观察中等阶层的白人家庭、工人阶层的白人家庭和贫困的黑人家庭在一些人类学指标上的表现。

尽管如此，我们可以推测，是文化使犹太人的表型智商得到了提高。（这种智商是环境使个人的基因构成发生了某种特殊的变化而形成的。）让我们回忆一下，1966年那个班级中的亚裔美国人，他们本身的智力属于平均水平甚至是略低于平均水平，但后来他们在职业中取得了极高的地位。他们的子女在很小的时候平均智商就达到了109分。那时这些孩子还没有进入美国的公立学校念书，和社区里其他非亚裔孩子也没有过多的接触。当童年期快要结束时，这些孩子的平均智商就下降到了103分。但是1966年那个班级中的亚裔美国人的孙辈，一定比他们的父母具备更多的智力优势，因为他们父母的智商比他们祖辈的智商更高，因此也就有可能创造出更有利的环境，结果就是1966年那个班级中的亚裔美国人的孙辈，其平均智商或许比103分还要高。

亚裔美国人的例子，即一代人会因为上一代人取得了较高的社会经济地位而智商有所提高，说明了文化是造成犹太人与非犹太人之间日积月累的智力差距的重要原因——也许是全部原因。

总之，在犹太人取得的智力成就中，有可能至少有一部分是超预期的成就，因为这样辉煌的成就远远大于根据其智力优势预测出的成就。

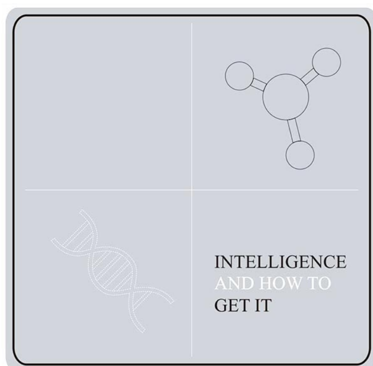
---

1. 顺便提一句，不考虑智商为140分的概率是高还是低，假设犹太人的平均智商是110分，那么如果随机选取一个犹太人，其智商高于非犹太白人的平均智商的概率只有75%；假设犹太人的平均智商是115分，那么这一概率就会达到84%。这就是将正态分布曲线按照与平均值的不同距离向上平移得出的惊人结果。

2. 在结束有关犹太人智商的话题之前，我应该提醒大家注意，犹太人的智力水平存在一个反常的情况。如果在美国人中随机选取一个包含大量犹太人的样本，测试结果显示，虽然犹太人的言语智商和数学智商比非犹太人的平均水平高出10～15分，但在关于空间关系能力（在大脑中对物体进行二维和三维空间操作）的测试中，他们的得分比非犹太人的平均水平低10分左右。这绝对是非常大的差异，我不知道还有哪个民族与犹太民族一样，在不同测试中的得分相差20～25分之多。不过，我绝对相信这一差距是真实存在的。我认识大概六七个犹太人，他们在各自的领域中都是顶尖的人才，但当走出一家餐馆时，他们却常常搞不清楚方向。我认为唯一与基因有关的种族差异，就是犹太人在空间推理能力上的不足。为何如此，我无法解释，但是我注意到，这似乎能解释为什么犹太人在沙漠里游荡了40年之久！

3. 查尔斯·莫瑞在其著作《人类的成就》中记述了不同欧洲国家的成就巅峰。

第10章
提高子女的智力，还有你自己的！



在这一章中，我会带着大家温习一下我们已经了解到的提高子女及自己智力的方法，然后再告诉大家一些你们可能相信但却没有什么证据能够证明的方式，以及一些令你们惊奇的提高智力的方法。我所说的智力，是指解决问题的能力和推理能力，虽然智力测试的分数和学业成就并不能完全反映一个人的智力水平，但我们在这里还是用它们来衡量智力的高低。

家长能用的最显而易见的方法

首先，有时无须有意识地去考虑采取何种策略，家长们一些本能的行为就能够提高孩子的智力：使用丰富的词汇和孩子交谈；让孩子参与成年人的谈话；念书给孩子听；对孩子尽量少一些责备，多一些鼓励；鼓励他们从周边的环境探索新知。避免过度压力，虽然大家都会这样做，但可能并不是出于本书中提到的原因：压力会损害边缘叶和额皮层之间的路径，从而影响人的学习能力和解决新问题的能力，极端的压力还会损伤人的记忆力。

教孩子如何将事物分类并在事物之间进行比较。鼓励孩子分析并评估这个世界一些有趣的方面。在放学后和暑假中给孩子提供能够启发智力的活动。（不过我要承认，我觉得有些父母给子女安排的活动多得有点儿过头了。有的家长把孩子的日程表排得满满的，刚参加完曲棍球训练就去上钢琴课，或者参加幼年童子军的各项活动，这使得孩子筋疲力尽，这种做法我个人并不推荐。）试着引导孩子多接触能够激发学习兴趣的同龄人。

上述这些都是社会经济地位较高的父母更常采用的方法，所有这些方法都与孩子的智力水平相关。我们承认，虽然从大体上来说这种相关性是存在的，但是我们并不清楚究竟是这样的活动能够提高孩子的智力，还是说这些活动只是更聪明的家长会对子女采取的行动（使这些孩子的智商高于平均水平的是父母的优良基因，而非养育子女的模范行为）。另一方面，常识也告诉我们，这样的行为应该对智力发展有利，至少总不会有什么害处。而且我们知道，穷孩子如果被社会经济地位较高的父母养大，智商和学业表现都会更出众。

令人半信半疑的益智方法

虽然你可能在新闻里看到一些关于提高智力的方法的报道，但是这些方法却没有多少效果——至少没有什么证据能够证明它们的成效。“小小爱因斯坦”（Baby Einstein）系列益智玩具可以四处移动并和孩子进行交流，这虽然有可能吸引孩子去探索，但也同样有可能造成孩子一味被动地接受。没有证据显示让孩子——不管是已出生的孩子还是尚在腹中的胎儿——听莫扎特的音乐能够提高智力。一些研究显示，在婴幼儿早期进行更多启发智力的活动会促进神经元的生长，从而提高孩子解决问题的能力。但这些都是以动物为对象的研究，研究人员将受到最少的智力启发的老鼠（关在笼子里，四周漆黑一片）和能够在有趣的环境中和同伴随意玩耍的老鼠进行比较。我们在一些婴儿身上也发现了这种智力的提高。这些婴儿原本几乎没有受到任何的智力启发，他们在正常的环境中成长之后，智力却有了大幅提高。我们不知道诸如使用别致的玩具等高水准的智力启发方式，能否提高人类婴幼儿的智力水平。

不过似乎还有其他很多方法能够提高智力水平，并且我们有充分的证据能够证明，其中大部分方法都是有效的。

与身体有关的方法：体育锻炼和母乳喂养

出生时体重较重的孩子比出生时体重较轻的孩子长大后更聪明。我们不知道出生时的体重较重就代表着更高的智商，还是只与高智商具有一定的相关度。不过既然有些方法能够增加孩子出生时的体重，为什么不试一试呢？我要说的方法就是体育锻炼。女性每天在跑步机上运动20分钟，每周坚持几次，就会增加孩子出生时的体重。体重较重的孩子当然也就更健康，加之大脑容积与体重有关，这样的孩子或许长大以后会更聪明。经常锻炼的女性生出的孩子脑容积更大。我们知道，一般来说，脑容积较大的人更聪明。

锻炼不但对小宝贝和准妈妈们有益，还能使每个人都受益。锻炼大肌肉群

实际上会促进神经元的生长，至少进行锻炼可以增加对大脑的供血量。即使在很大年纪才开始锻炼也对智力有好处。实验表明，和不经常锻炼的老年人相比，经常锻炼的老年人解决问题的能力能够保持得更久。所有的研究都显示，每天进行30分钟以上有关液态智力的锻炼，能够取得0.5个标准差的效果。力量训练加上对心脏的锻炼，要好过只进行心脏锻炼的效果。在中年能够定期锻炼的人在70多岁时患阿尔茨海默病的概率，比不定期锻炼的人低1/3。即使是60多岁才开始锻炼，也可以将患阿尔茨海默病的概率降低一半。

为了提高孩子的智力，还有一件可能非常重要的事情是妈妈们能够做到的，那就是母乳喂养。如果对遗传基因极普通的孩子进行9个月的母乳喂养，就有可能使孩子的智商提高6分（而母乳喂养超过9个月似乎没有什么益处）。母乳喂养对于早产儿尤为重要。

提高液态智力的练习

有几种活动能够提高液态智力，而且适用的人群不仅仅局限于儿童。我们来回忆一下，液态智力是指解决新问题的能力，之前学习到的规律和概念不一定对解决这种问题有所帮助。最典型的例子就是瑞文推理测验。你在测试中会看到不同的几何图案，并且这些图案按照某些特别的方式发生了变化，所以你要做的就是当即找出这些图案发生变化的规律，并且根据这一规律判断这些图形将要发生的变化。能够提高液态智力的活动包括锻炼注意力控制和工作记忆能力的电子游戏。

神经科学家萝莎里奥·罗达及其同事介绍了几种能够提高儿童液态智力的练习，其中一项是预测练习。他们使用的一项练习就是教孩子们预测一只潜入池塘中的鸭子会从什么地方浮出水面。孩子们使用操纵杆控制一只猫，守在他们认为鸭子会浮出来的位置。另外一项练习与刺激辨别有关。孩子们要记住一些卡通图像的特征，然后从一列卡通图像中把这个卡通图像挑出来。还有在第3章中提到的冲突解决练习和自我控制练习。这些练习都可以改善在瑞文推理测验中的表现，因为这项测验测试的就是注意力控制和工作记忆。这些与液态智力有关的练习对孩子们在青春期之前的学习来说是非常重要的。

儿童神经学家托克尔·克林伯格及其同事发现，各种锻炼工作记忆和注意力控制的任务能够帮助患有注意缺陷多动障碍的儿童改善注意力。其中一些任务无须借助电脑就可以完成。例如，实验人员向孩子们读出一组个位数（例如4、7、2、9、5），然后要求孩子们按照相反的顺序说出这些数字。还有一些任务是在电脑上进行的。例如“动—不动”（“go-no go”）的练习，孩子们会看到两个灰色的圆圈，如果其中的一个圆圈变成了绿色，

孩子们就要按一个键；如果变成红色了，则不按键。这样的练习不但可以提高孩子们的反应速度，还可以降低孩子们在未受过提前训练的任务中出现错误的概率，并提高在瑞文推理测验中的分数。所以接受过这些锻炼的孩子在瑞文推理测验中的分数，都超过了没有接受过这种锻炼的孩子的平均分。类似的练习还可以改善普通的健康成年人的工作记忆以及在瑞文推理测验中的表现。

最后，如果连续5天进行中医保养中经常使用的冥想练习（呼吸练习、姿势练习和身体意识），就能够改善执行功能，并提高在瑞文推理测验中的分数。如果连你们都对这种练习将信将疑的话，那么我就更不可能相信了。但是这项研究的作者中包括唐一原和迈克尔·波斯纳这样极受人尊敬的神经科学家，所以我信了。

提高孩子的自我控制能力

我们掌握的最有力证据显示，自我控制能力较强的孩子智商更高，并且不论智商高低，这些孩子的学习成绩都更好。人格心理学家沃尔特·米歇尔及其同事发现，在斯坦福大学幼儿园里大部分来自中上阶层家庭的孩子中，那些为了获得更好的回报（两块饼干）能够将眼前的满足感（一块饼干）延后的孩子到了十几岁，学习成绩更优秀，在SAT考试中的分数也远远高于其他人。纽约市的那些社会经济地位较低的少数族裔孩子，如果学会延迟满足感，也会得到更高的分数。虽然延迟满足感的能力和未来的能力测试中的高分数之间存在一定的联系，但是我们并不清楚，这是否仅仅是因为拥有较高智力的孩子在4岁时恰好拥有更强的延迟满足感的能力。他们长大以后变得更聪明，不是因为延迟满足感的能力有助于学习，而是因为由于遗传或在其他环境因素的作用下，他们注定比别人更聪明。不过，似乎延迟满足感的能力本身就会提高人的能力，因为更强的自我控制力能够使学习变得容易一些。让我们回想一下，心理学家安吉拉·达克沃斯和马丁·塞利格曼的研究发现，在费城一所磁铁学校的初中生中，拥有更强自我控制力的平均学分绩点（GPA）也更高。实际上，自我控制力和平均分数之间的关联度，是智力与平均分数之间的关联度的两倍。在这一因果关系上我们拥有更加强有力的证据。我们几乎能够肯定，自我控制力能够使一个人取得超出其智力水平的成就。

不幸的是，我们并不太清楚哪些方法可以提高孩子们的自我控制能力，不过研究还是给我们留下了一点提示。我们可以确认的是，如果大人们没有取得什么成绩就对自己进行奖励，那么孩子们很有可能有样学样。但是如果他们看到大人们在取得了很大的成就以后才犒劳自己，孩子们就会以大人们为榜样。同样，米歇尔及其同事还发现，一些小技巧可以帮助研究中的孩子们为了更大的回报而延迟立即得到好东西的满足感。当研究人员让

孩子们“思考一些有趣的想法”而不是思考即将得到的回报时，孩子们能拖延得更久。当研究人员鼓励孩子们忘掉回报，把回报放到他们视线以外，孩子们也能拖延得更久。我们不知道这样的建议是否适用于实验以外的行为，或许是可以的。如果家长要寻找合适的机会培养孩子的耐心，尤其是当他们向孩子提出如何培养耐心的建议时，这样的建议或许是有效的。家长们或许可以试试给孩子们做个延迟满足感的榜样。米歇尔及其同事发现，孩子们会模仿成年人的行为。有些孩子会看到大人马上接受了眼前的奖励，而并没有等待得到更大的奖励。大人可能会这么解释：“你可能注意到了，我是一个喜欢享受现在的人。要是在生活里什么事情都等来等去的，那永远也不能真正地生活。”那么即使孩子们本来想要延迟满足感的，看见这样一个榜样后，大多数情况下也会马上接受别人给予的奖励。

夸孩子聪明伶俐不如夸孩子刻苦努力

家长一定要告诉自己的子女，他们的智力在他们自己的掌控之中。亚洲人尤其相信，只有努力付出才能够提高智力。亚裔美国人比欧裔美国人更加努力，所以实现了更高的学习目标，这一点也不令人吃惊。而且亚洲人在遭遇失败后会比获得成功后更加努力，不像欧裔北美人在成功后比失败后更加努力。如果不成功，就要更努力——这一点家长们一定要言传身教给孩子。

夸奖孩子聪明恐怕不是一个好主意。相反，应该表扬孩子的艰苦努力，因为这是他们可以直接控制的。夸奖孩子聪明的不妥之处就在于，这会使孩子一心只想要表现自己有多聪明，所以他们就只完成自己擅长的任务，而不会努力去自己改善自身存在的不足。换句话说，如果夸奖孩子聪明，他们就会拒绝接受挑战，拒绝完成可以从中学到很多东西的任务。

有一项非常巧妙的实验验证了这一点。发展心理学家克劳迪娅·米勒和卡罗尔·德韦克告诉一群孩子，他们很好地完成了瑞文推理测验中的试题，然后表扬其中一部分孩子很聪明，表扬另外一部分孩子很努力。随后研究人员让孩子们完成另外一组试题——这些试题有的比较容易（“这样我的表现就会很好”）有的比较难（“这样我就能够从中学到很多东西，即使这会使我看起来不那么聪明”）。被称赞聪明的孩子中有66%选择了能够表现他们很聪明的容易的试题，而被称赞努力的孩子中有超过90%的孩子选择了能够从中学到东西的试题。如果孩子们的成绩好是因为他们聪明，他们就不愿意冒险去发现他们其实并没有那么聪明。如果孩子们的成绩好是因为他们努力，他们就愿意找一些更难的题目挑战自己的极限，并从中学到更多的知识。

实际上，在孩子们开始选择不同难度的试题之前，米勒和德韦克还要求孩

子们完成另外一组试题，难度比第一组试题大很多。然后研究人员让孩子们解释造成他们在这组试题中表现欠佳的原因。被称赞聪明的孩子往往认为是因为自己能力不足，而被称赞努力的孩子则认为是因为自己努力不够。被称赞聪明的孩子大多不愿意继续完成这组试题，而且也不如被称赞努力的孩子那样投入。作为附加的实验，米勒和德韦克让孩子们完成第三组试题，最初被称赞聪明的孩子完成的题目，少于被称赞努力的孩子。这则实验的寓意很清楚：不要夸奖孩子聪明，要夸奖孩子努力。

不要轻易向孩子许诺奖励

如果孩子对你想要鼓励他（她）做的事情已经有了一些兴趣，就不要再向孩子许诺任何奖励。我和发展心理学家马克·莱珀、戴维·格林一起观察了一群幼儿园的孩子参加一项新颖的活动——用神奇笔画画。大多数孩子都用这种笔画画，显然他们玩得很开心。我们向一些孩子许诺，如果他们用神奇笔为我们画一些画，就可以得到一个奖励。这些孩子愉快地答应并画了画。一两周后，我们又把神奇笔发给孩子们画画。那些因为用神奇笔为我们画画而得到奖励的孩子，比没有得到奖励的孩子画的数量少得多——画的质量也大大下降。实际上，“契约”已经把游戏变成了工作。我们表扬了另外一些没有得到许诺的孩子画得很好，随后这些孩子用神奇笔画出的画，比既没有得到许诺也没有得到表扬的孩子要画得多。所以如果你想让孩子做什么事情，一旦他们做了，就表扬他们，不要向他们许诺如果做了就会得到奖励。

不过有时候向孩子许诺奖励也是一个不错的主意。如果孩子没有得到外部奖励就不会去做某件事，那么许诺一定的奖励才是适当的做法。如果孩子一开始对一项活动缺乏兴趣，那么奖励也许能够吸引孩子去尝试一下，没准孩子最后会真正被这项活动所吸引。我认为KIPP学校采取的奖励机制也许对于学生来说就是个很好的主意，而这些学生之前就读的学校，并没有采用奖励制度来激发学生的兴趣。

有效的辅导

当辅导孩子学习时，尽量记住第4章中提到的马克·莱珀总结的辅导5C原则：培养一种控制感（Control），向孩子发起挑战（Challenge），向孩子灌输信心（Confidence），培养孩子的好奇心（Curiosity），将问题放在真实世界的环境里或电影电视节目中出现的环境里，让孩子将问题放到一定的情境中思考（Contextualize）。不仅如此，不要总盯着一些小错误不放，例如忘记写负号等；尽量避免让孩子犯错，除非孩子可以通过这个错误吸取一些教训；不要因为要顾及孩子的面子而降低辅导材料的难度，相反，换一种方式表述这些材料。提出一些诱导性的问题；不要过多地表

扬孩子，以避免使孩子感觉自己总是在被评估。

选择一所好的学校

最后——还有一些有关学校的建议。在你力所能及的范围内，尽量让孩子进入由最好的老师负责的班级，尤其是孩子一年级的时候，避免进入新教师负责的班级。如果学校在阅读、数学和科学教学中还没有使用已被证明有效的电脑程序，试着让学校考虑一下这样做。访问美国教育部网站上的有效教育策略资料中心，你可以援引其中的一些章节，这些章节说明了哪些电脑程序是适合你的子女目前的成绩水平的。如果子女所在的学校没有采用任何合作学习的方法，也就是孩子们一起学习如何解决问题并拓展新的知识领域，那就鼓励学校这样做，也可以引用有效教育策略资料中心上的介绍。看一看学校的校长是否清楚学校里究竟哪些教师才是好教师，是否有可能对这些高水平的教师进行奖励。如果不可能对教师的素质进行奖励，那就敦促教育委员会使之成为可能。（工会合同也许不允许将教师资历以外的任何方面作为奖励的依据。如果是这样的话，你可以建议教育委员会奖励表现优异的学校中的全体教师。）说服教育委员会不要过于强调教师的资格和高学历，因为没有证据显示持有各种资格证书和高学历的老师就一定是会教书的好老师。应该花更多的时间培训教师的教学技能，可以让同龄教师或专家们在教室里旁听，然后给予反馈意见，帮助这些老师提高教学技能。

总之，你可以使用本书中提到的很多方法来提高孩子的智力——还有你自己的。

结语
关于智力和学业成就，我们都了解哪些内容？

坚定的遗传决定论者们认为，智力主要是由基因决定的。你的基因决定你有多聪明，你就只能有多聪明，没有什么环境因素——父母的养育方法，或者你就读学校的好坏——能够改变智力。即使不是全部，美国绝大多数研究智力的科学家和普通大众也都这样认为。这样的想法对于个人来说是非常不幸的，因为这就意味着即使努力付出也不能使“真正的”智力有所提高。这对于公共政策来说更是一场灾难，因为这意味着种种教育干预措施注定是要失败的。不过幸运的是，这种认为智力由基因决定的观点是错误的，现在我们就来看一下我们是如何发现这一点的。

智力的遗传度没有固定的数值。在某种环境中生活的一群人，和在另外一种环境中生活的一群人，两者智力遗传度的高低会有很大差异。如果生活的环境有利于智力的发展，那么智力的遗传度就确实要高一些——或许能够达到70%。发达国家里中上阶层的家庭就属于这种情况。他们提供的环境能够促进智力发展，在这方面每个家庭之间没有太大的差别。如果某个特定人群中的每一个人，生长的环境几乎完全相同，那么能够引起智力差异的唯一因素就是基因。中上阶层和这种情况比较接近，所以这个群体的遗传度就是决定智力差异的重要因素。

但是如果环境差异很大——不同家庭之间的环境差异很大——那么引起个体智力差异的主要因素就是环境。穷人就属于这种情况。对于穷人来说，只有10%的智力差异是由基因引起的，这也就意味着，改善穷孩子所处的家庭环境会对孩子的智力起到很大的作用。实际上，如果把贫困家庭的子女放到中上阶层家庭里抚养长大，那就有可能使其智商提高12分，甚至是18分之多。这对于学业成就的影响也很大——至少是0.5个标准差，在某些情况下甚至可能达到1个标准差。

虽然不同人群的智力遗传度高低不同，但遗传度不会限制可变度的大小——对于任何人来说都是这样。发达国家的人的身高在最近几代人里实现了很大的增长——这种增长和遗传基因没有任何关系。

在20世纪的100年中，人的智力水平也经历了极为相似的增长。在过去的60年里，智力测试的分数增加了超过18分，在过去100年里甚至增加了将

近2个标准差（30分）。不仅如此，在瑞文推理测验（几十年来都被认为测量的是与文化无关的真正的智力）中的分数在不到60年的时间里提高了2个标准差。

为什么会出现这种增长呢？根本原因很简单：学校和文化都发生了根本性的变化，这些变化会提高智力测试中很多小测验的分数。家长和学校越来越多地教孩子如何使用适合科学分析的分类方法对事物进行分类。孩子们也从媒体上了解了世界运转的方式——为什么警察要穿警服，为什么街上的地址要按顺序编号，为什么人们要纳税——这使得他们在智力测试理解小测验中的分数大大提高。瑞文推理测验——以及这一测验所测试的液态智力——的分数有所提高，至少有一部分原因是，在最近的几十年里，学校在教授算术时越来越多地运用几何和分析的方法，还有一部分原因可能是电脑游戏的作用。几年前，麦当劳的开心乐园餐中包含的迷宫游戏的难度，甚至超过了天才儿童参加的智力测试。

不仅如此，和以前的时代相比，人们接受的教育也越来越多了。在过去的100年里，人们接受学校教育的平均年限由7年增加到14年。既然在智力测试的分数中，一年的学校教育能够抵得上年龄增加两岁，那么如果智力在那段时间里没有发生极大的增长就是不正常的了。

智力增加的部分中有多少是属于“真正的”智力的增加？在这一点上我有一些话要说。首先，如果我们将智力广泛地定义为“推理、计划、解决问题、抽象思考、理解复杂概念……的能力”，那么1910年时10岁的人，他（她）的曾孙辈的智力不可能比他（她）高出两个标准差。不过，现在的人们确实比他们祖辈们的智力更高。在某种意义上，能够告诉你为什么街上的地址都要按顺序编号的孩子，就是要比不能解释清楚的孩子更聪明。能够运用归类法的术语寻找事物之间相似性，确实是一个很大的优势。推理的步骤，例如对假设进行验证的步骤，是每一层级的学校都会涉及的教学内容，这样的学习还可以解决日常生活中出现的问题。通过广泛教授有关概率论和成本收益推理的知识，“计划”和“选择”这两个智力组成部分的能力也都得到了提高。

既然学校可以使孩子们变聪明，我们就无须怀疑好学校能使孩子们变得更聪明。虽然教育券、特许学校、全面的学校干预措施、教师的资格和高学历都不一定会提高教育的质量，但是其他一些因素可以——有的还能起到很大的作用。教师素质的高低有很大差异，所以想办法提高教师的素质和能力能够产生很大的效果。如果每年我们能用平均水平的教师淘汰表现最差的5%的教师，那么短短几年内孩子们的学习成绩就会取得突飞猛进的进步。使用电脑辅助教学也会大大提高学习的效果，有些合作学习的方法也是非常有效的。回忆一下赫恩斯坦在委内瑞拉进行的强化项目。这一项目不仅大幅提高了普通中学生解决问题的能力，还使他们的智力实现了不

小的增长——在测试多重问题解决能力的测验中，其得分提高了5分。

对于社会阶层和智力的关系，大家普遍认为，主要来自于遗传的智力决定了一个人的社会阶层。聪明的人遗传基因更优秀，所以他们注定要进入社会的上层，而不聪明的人遗传基因较差，所以注定要落入社会的底层。的确，智力有一部分是来自遗传，而且平均来说，更聪明的人因为继承了高智商的基因，所以一般都处于较高的社会阶层。不过我认为遗传基因对于决定一个人的社会阶层只能起到很小的作用。社会经济状况最差的1/3的孩子和社会经济状况最佳的1/3的孩子，其平均智商相差大约10分。我们知道造成这一差距的一部分原因是生物原因而非遗传因素，包括锻炼，母乳喂养，母亲在怀孕期间饮酒或吸烟，以及接触各种危险的化学品和污染。还有一部分原因是社会经济地位较低的孩子就读的学校水平较低，他们接触的同龄人大都会使他们的智力水平有所降低。我们还知道，社会经济地位较低的家庭对子女进行社会化的方式，既不能促进他们的智力发展，也不能让孩子做好上学的准备。不仅如此，如果将出生在社会经济地位最低的1/6家庭中的孩子，交给社会经济地位最高的1/4的家庭抚养，这些孩子的智商会提高12～18分，所以这并没有给基因在社会阶层的方程式中留下什么空间。我并不怀疑基因会起到一定的作用，但我很难相信遗传基因会在很大程度上决定人们所处的社会阶层。社会经济状况最差的1/3的孩子和社会经济状况最佳的1/3的孩子，平均智商相差了10分，这样的差距绝大部分——如果不是全部——从根本上说是由环境因素造成的。

我们能够满怀信心地说，种族之间在智力上的差异与遗传没有任何关系。绝大多数证明种族之间智力差异是遗传因素造成的证据都是间接证据，很容易就被驳倒了。美国的黑人中，有的全部都是非洲血统，有的大部分是欧洲血统，我们从自然实验中得出的几乎所有直接证据都显示，智力与遗传基因的差异没有任何关系。而且种族之间在智力和学业成就方面的差距正在以每代人大约1/3个标准差的速度缩小。黑人现在的平均智商就高于1950年时白人的平均智商。

《不让一个孩子落后法案》要求学校在半代人的时间里消除不同社会阶层和种族的学生在学习成绩上的差距，这是很荒谬的。这忽略了一个事实，那就是社会阶层和种族的差异在婴儿期初期就存在了，而且这些差距不仅与学校有关，还与经济因素、居住的社区和文化差异都有关系。

这对于消除差距来说实在是个坏消息。不过好消息就是，社会经济地位较低的孩子和少数族裔的孩子是有可能提高智力和学习成绩的。我们至少已经能看到这种提高的雏形了。我们已经尝试过一些不彻底的办法了，但没有取得很好的效果。我们需要对穷孩子进行高强度的儿童早期教育，我们需要进行家访，教父母如何促进孩子的智力发展。这样的努力不仅能在短时间内使孩子的智力得到很大提高，还能使他们在学习成绩和职业成就方

面取得长远的进步。大规模的小学、初中和高中教育计划也能够使孩子们获得更高的学业成就。还有一些简单又不用花钱的干预措施，效果最明显的就是使学生相信他们能在很大程度上掌控自己的智力，这能够使学生们显著提高学习成就。

相信自己可以掌控自己的智力——再加上看重成就的父母——就能够创造出奇迹。至少亚洲人和犹太人验证了这一点。没有可靠的证据能够显示，具有东亚血统的人和具有欧洲血统的人在智力遗传基因上有什么分别。实际上，智力测试显示两者的智力不相上下。甚至有一些证据显示，东亚孩子刚上学时的智力还不如美国的白人孩子，但是经过几年的学习以后，这种差异似乎消失了。但是东亚人的学业成就——尤其是需要付出努力的数学和科学——比欧裔美国人不知要高出多少倍。具有东亚血统的美国人与欧裔美国人在智力上也没有多少差别。此外，亚裔美国人所取得的学业成就和职业成就，远远超过了他们的智力水平预示的他们“应该”取得的成就。亚洲人和西方人之间存在差距的原因就在于亚洲人刻苦努力并且能够持之以恒。

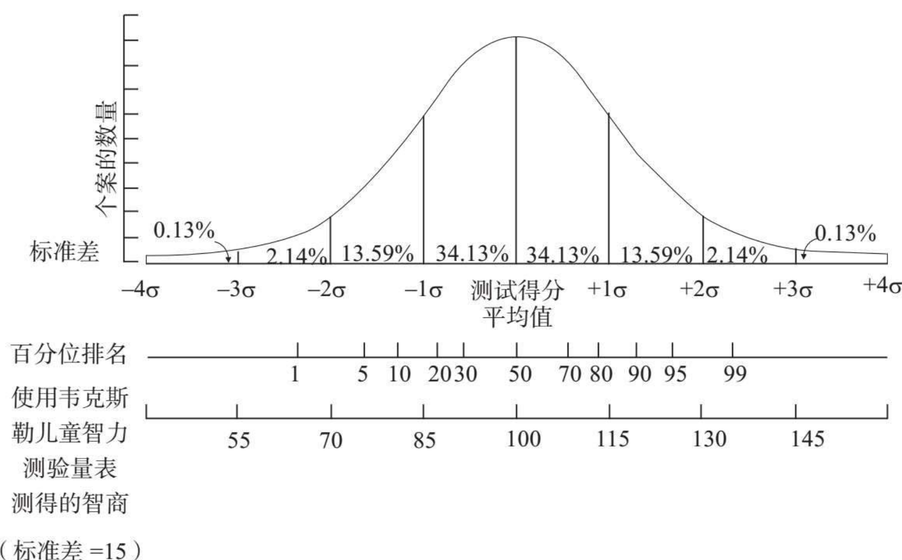
犹太文化毫无疑问也取得了同样有益的成效。一般来说，犹太人看重成就，尤其是智力成就。犹太人和非犹太人所取得的最高的智力成就之间，存在着很大的差距，而这些差距不是遗传基因造成的。中世纪时欧洲人与阿拉伯人和中国人之间存在着更大的差距，从中世纪（意大利和英格兰交替处于领先地位，苏格兰则在不到200年的时间里，从未开化的蛮荒状态发展到精明睿智的文明状态）开始，欧洲各国在不同历史时期也存在差距，美国各个地区之间也存在差距，这些都不是基因造成的。我们要面对的是犹太人与非犹太人在智力上存在2/3个标准差的现实，这个差距至少有一部分是文化原因造成的。

最后，有很多种方法可以帮助我们提高孩子和我们自己的智力和学业成就。不论是生物方面的方法（体育锻炼，孕妇避免吸烟、饮酒，对新生儿采用母乳喂养）还是教育方法（教会孩子如何进行分类，遵循有效的辅导原则），都能够大大提高智力水平。

现在我们终于可以卸下所有遗传决定智力的思想包袱了。相信智力在我们自己的掌控之中，这种信念本身虽然并不会让我们变得更聪明，但它是个很好的开始。

附录一
对一些统计术语的非正式定义

我们经过研究发现，各种各样的现象都呈正态分布，就像图A–1中显示的那种钟形曲线。例如，如果我们要画图显示不同母鸡每周产蛋的数量，生产某型号汽车的过程中出现错误的数量，或者一群人的智力测试分数，代表这些数据的曲线的形状大体上都接近钟形。我们无须了解这些分布呈现钟形背后的数学原理。重要的是正态分布曲线用处很大，能够帮助我们找出一个观测值与另一个观测值之间的联系。图A–1中显示的正态分布曲线按照标准差——用“标准差”来代表平均分与平均值的偏离程度——分为几段。如果观测值的数量足够多，正态分布的曲线就无限接近钟形，也就是非常标准的正态分布。在标准的正态分布中，大约68%的观测值都分布在偏离平均值（也就是图A–1中的曲线上0的位置）+1或–1个标准差的范围内。标准差的概念还有一些用处，那就是百分位排名与标准差的关系。在所有的观测值中，大约84%的数据都集中在+1个标准差或少于+1个标准差的范围内；超出平均值+1个标准差的观测值在整个分布中对应的百分位排名就是第84位。几乎98%的观测值都集中在+2个标准差的范围内。正好超出平均值+2个标准差的观测值在整个分布中对应的百分位排名就是第98位，剩下的2%观测值对应的百分位排名则高于第98位。几乎所有的观测值都位于+3或–3个标准差的范围内。按照惯例，大多数智力测试都将智商的标准差设定为15分（平均智商设定为100分）。



图A–1 正态分布曲线

标准差这一单位非常有用，我们可以用标准差来表示效果值的大小。例如，一项新的教学方法改善学习效果的程度就可以用标准差来表示。最常用的效果值的指标叫作Cohen’s d，是这样计算的：A组的平均值减去B组的平均值再除以A、B两组的标准差之和（或者有时候只除以A组的标准差）。

按照惯例，d的值为0.2或更小就被视为效果甚微。这相当于将实验组分数的百分位排名从第50位提高到将近第60位。如果新的教学方法（百分位排名第60位）和老的教学方法（百分位排名第50位）相比能够使孩子的成绩排名提前10名，你或许并不认为这是很小的效果。你是否愿意付钱使用新的教学方法，一部分也取决于百分位排名从第50位上升到第60位究竟具有多么重要的意义。如果你用孩子盲打的速度达到每分钟40字的熟练程度所需要的时间来衡量教学的效果，只需要几天的时间就可以将百分位排名从第50位提高到第60位，很可能你就不愿意为这样的进步花费太多的钱，也不愿意学校在上面花费太多。如果你根据SAT数学考试的平均成绩来衡量两所高中采用的数学教学方法的效果，采用一种教学方法后数学考试的平均成绩是500分，而采用另一种教学方法的平均成绩是520分，这就是百分位排名第50位与第60位之间的差距（假设SAT分数的标准差是100）。你或许愿意为了孩子的分数得到这样的提高而不惜花上大笔钱财。或许你还乐意你们的教育委员会为每位学生花上一些钱，以采取更有效的教学方法。

按照惯例，d的值为0.5左右被视为是一般的效果。不过，在智力测试和学业成就的概念里，这么大的效果值一般已经相当引人注目了，这相当于SAT数学部分的分数从500分提高到550分——有时这样的差距就是考取中上水平的大学和考取名牌大学的差距。为了使普通孩子的SAT数学成绩百分位排名从第50位提高到大约第70位（这就是0.5个标准差对应的效果值），你和学校或许愿意花大价钱采用新的教学方法。

0.7~1个标准差被认为是很大的效果。1个标准差对于教育和智力的差距来说是非常大的。一般认为黑人与白人的智商差距就接近1个标准差。在第6章中，我们讨论了黑人与白人的实际差距是否有如此之大。如果有，那就意味着黑人的平均智商在白人的智商分布上对应的百分位排名是第16位。如果一项干预措施能够使孩子在全国数学考试中的百分位排名从第50位提高到第84位，那么人们一定认为在这项干预措施上花一大笔钱是值得的。对于一个国家来说，如果数学成绩取得了这样的进步从而使国家的竞争力有所增强，那么国家就应该不惜成本地大力普及这项干预措施。

相关系数可以用来测量两个变量之间线性联系的程度。例如，智商与学习成绩之间的相关度恰好是0.5左右，也就是说两者的联系较强。不过智商与学习成绩之间至少应该具备相当的关联度，因为设计智力测试就是为了预测人们在学校内的学习成绩。相关系数的变化范围是–1到＋1之间，–1表示两个变量完全成反比，＋1表示两个变量完全成正比。相关系数为0就意味着两个变量之间没有任何联系。相关系数也可以用来衡量效果值的大小，或者关联度的大小。相关系数的数值小于0.3就是较小的相关度，0.3~0.5就是中等大小的相关度，在0.5以上就是较大的相关度。但是，和效果值一样，关联度是否重要取决于关联度中的变量，而不是相关度的大小。我们也可以用标准差的概念来解释相关系数。如果两个变量的相关度是0.25，那么第一个变量增加1个标准差，第二个变量就会增加0.25个标准差；如果相关度是0.5，那么第二个变量就会增加0.5个标准差。如果班级的规模与学生在标准化考试中的成绩之间的相关度是–0.25，那么班级规模缩小1个标准差，学生的考试成绩就应该提高0.25个标准差（假设班级规模与考试成绩之间确实存在因果关系）。

多元回归是一种分析方法，它能够找出一些因变量（或预测变量）与结果变量（或某个目标变量）之间的关系。例如，我们或许想要比较哪些变量能够最大限度地预测一所房子在房地产市场上的吸引力。我们可能会衡量房屋的面积有多少平方米，有几间卧室，卫生间的舒适程度（例如水池的数量、是否能洗热水澡、使用的材料质量好坏等），小区的平均收入，以及由潜在买主的数量代表的房屋的抢手程度。然后我们将这些变量同时与房屋在市场上的吸引力联系在一起，也就是房屋在市场上的售价——目标变量。不考虑所有其他变量对房屋价值的贡献度（即将所有其他变量设为

常数）的条件下，找出某个变量与市场价值之间相关度的大小，就可以估计出这个变量对于房屋市场价值的贡献度。因此，当将所有其他变量设为常数时，房屋的抢手程度与市场价值之间的相关度可能是0.25，卫生间的舒适程度与市场价值之间的相关度可能是0.1。不过所有这些变量之间都存在一定的相关度，并且对某些变量的测量可能比其他变量更为准确，某些变量可能与其他一些变量存在一定的因果关系，而与另一些变量不存在因果关系，有些没有测量的变量或许会对一些得到测量的变量产生一定的影响。这就导致多元回归的结果有可能对我们产生误导。房屋的抢手程度与市场价值之间的实际相关度，可能远远高于多元回归分析得出的0.25，也可能远远低于0.25。

有无数个例子能够说明，多元回归分析得出的因果关系，往往与实验得出的因果关系不一致。从因果推论的角度来看，这些实验几乎都比多元回归分析更加可取。例如，大概在15年前，我曾经参加过全美卫生研究院召开的共识发展会议。这次会议的目的就是要重新审视关于冠状动脉阻滞的治疗方法的研究，究竟是临床治疗还是手术治疗效果更好，并就这两种治疗方法的适当性达成共识。要审查的研究中有大量研究都是由政府资助的，耗资巨大。在这些研究中，研究人员将大量和患者有关的变量放入一个多元回归方程式中，例如病史、年龄以及社会经济地位等，然后在不考虑不同患者使用的其他治疗方法的情况下，确定某种治疗方法的疗效。但是由于美国管理研究政策的内部审查委员会要求，必须给予患者选择治疗方法的自由（不过我们并不能确定这样做实际上是否真的符合患者的利益），所以所有的实验证据都因为带有自我选择的人为因素而动摇了。不过除了美国的研究以外，还有两项欧洲的研究，都达到了随机为患者安排各种治疗方法的标准。因此，专门小组的成员放弃了耗资巨大的美国研究，只对两项欧洲研究的结果进行了分析。

让我们再考虑一个与本书更为相关的例子，那就是班级规模是否真的会影响学习成绩。多元回归分析告诉我们，不考虑学校的规模、学校所在社区全部家庭的平均收入、教师的薪酬、具备资格的教师的比例，以及学区内每名学生的教育支出等因素，规模一般的班级与学生的学习成绩之间没有联系。在另外一项随机选取研究对象并实施得很好的研究中，进行比较的班级学生人数相差很多（将有13～17名学生的班级与有22～25名学生的班级进行比较）。结果这项研究发现，将班级人数缩减至13～17人，学生们在标准化考试中的成绩提高了0.25个标准差以上——这对黑人孩子起到的作用也大于白人孩子（1999年）。这不仅仅是另外一项有关班级规模的影响的研究，它还取代了所有有关班级规模的多元回归研究。

在本书中我偶尔会提到多元回归研究，不过每次都会提醒大家注意研究的结果。

自我选择是相关性研究和多元回归分析面临的难题之一。有很多原因要求我们必须要对自我选择有所了解。如果我们说智力与未来的职业成就之间具有一定的相关度——比如0.4，那么大家往往就会条件反射地认为，这样的关系完全是因果关系——智商更高的人工作能力也更强。但是智力与其他因素也是相关的，例如智商较高的孩子通常其父母的社会经济地位也较高。如果父母的社会经济地位较高，那么不论这个孩子智商高低，都更有可能念大学。同样，不论这个孩子的智商高低，拥有大学学历就有可能获得更高的职业地位。因此，智力与职业成就之间的关联就受到了其他因素的影响，例如父母的社会经济地位以及大学学历，那么这个孩子，或是研究的对象，就具备了“自我选择”的自由。（说一个人“自我选择”了父母的社会经济地位恐怕并不恰当，因为显然这是不可能的。但是由于真正进行比较的研究人员不能决定这个变量的大小，所以就好像是研究对象决定了这个变量的大小。总之，一些与研究对象有关的变量是研究人员无法控制的，其大小也不能由他们选择，有时他们甚至不知道这些变量的变化。）

只要一项研究仅对一个特定的变量进行测量而不是控制，我们就一定要注意到，对测量的变量（以及所有其他测量或不测量的变量）的大小进行选择的是研究对象，而不是研究人员。这就在很大程度上限制了推理的准确性。在班级规模的例子中，研究人员使用了多元回归的方法，使得班级规模这个变量的大小可以进行自我选择（也就是说，研究人员不决定班级规模的大小），并且班级规模这个变量也许与其他各类变量都有联系，这些变量有可能会扩大或限制班级规模对于学习成绩的影响。能够完全避免自我选择问题的唯一方法，就是让研究人员来选择因变量或预测变量的大小（例如，是人数较多的班级还是人数较少的班级），然后观察因变量或预测变量对目标变量的效果（例如成就测试的成绩）。由于不是总能做到这一点，所以我们只好满足于相关性分析和多元回归分析，同时对自我选择的问题保持警戒。

最后，统计显著性告诉我们，如果一项计划实际上并没有任何效果，那么研究得出的效果值——例如班级规模对于学习成绩的效果——就有可能是偶然产生的。按照惯例，如果统计显著性的值为0.05，这就是说，在与进行的研究设计相同的研究中，两个平均数之间的差距或某个特定大小的相关度在100次统计中只有5次出现的概率，或者20次中只有一次出现的概率。不过统计显著性在很大程度上取决于观测值的数量。如果观测值数量足够多，那么即使在实践中或理论上小到不具备任何意义的差距，在统计学上也可能会很重要。我在本书中提到的任何一项研究所得出的结果，其统计显著性都达到了低于0.05的水平。只有一项我称为“不太重要”的研究结果，其发生的概率低于0.1，这项研究的结果有可能是偶然获得的。

附录二
黑人与白人的智力差异完全由环境所致的论据

在这个附录中，我会列举出证明黑人与白人的智力差异主要由基因所致的依据，并逐一进行驳斥。理查德·赫恩斯坦和查尔斯·莫瑞（1994年）合著的《钟形曲线》中有关种族和智力的一章，以及菲利普·拉什顿和阿瑟·詹森（2005年）最近的一篇文章中，都提出了遗传基因造成黑人与白人智力差异的论据。而且对于下列基因决定论的观点，不少科学家或多或少都抱着支持的态度。

1. 白人的智力遗传度很高。因此，有可能黑人与白人的智力差距中的一大部分都应该归结于遗传基因。

2. 有些智力测验宣称与文化无关，例如涉及对不同几何图形之间的关系进行分析的测试。有的测试与文化有很大关系，例如找出沙龙舞和方块舞之间的相似之处。黑人与白人在前一种测试中的差距，大于两者在后一种测试中的差距，所以这种差距不太可能是文化原因造成的。

3. 撒哈拉沙漠以南的黑人的智商（赫恩斯坦和莫瑞1994年的测量结果是75分，拉什顿和詹森2005年测量的结果是70分）比非裔美国人更低。非裔美国人的智商（一般认为是85分）比较高是因为他们的基因库里有大约20%的欧洲基因。

4. 黑人在所谓的g（一般智力的g因素）负荷量较高的试题上的表现，不如在g负荷量较低的试题上的表现，而在g负荷量较高的试题上的表现，比在g负荷量较低的试题上的表现更易受到遗传基因的影响。

5. 近交衰退（近亲结婚对后代智力造成的不利后果）会影响在某些智力测试题目上的表现。在这些题目上黑人与白人之间智力的差距最大，这也就表示黑人与白人在最易受基因影响的测试部分上的差距最大。

6. 黑人和白人的智商都与脑容积——大脑的体积——有关，白人的脑容积一般比黑人的大。

7. 智商更高的人的反应时间更短。在测试中，要求被测试者将手放在一个按钮上，当一个灯泡亮起时，尽可能快地触摸灯泡。白人在这个测试中的反应时间比黑人短。（反应时间短就意味着反应快。）

8. 由于受到基因的影响，黑人的平均智商天生就比白人低。如果黑人孩子的父母与白人孩子的父母智商相同，那么黑人孩子的平均智商会逐渐降低到白人孩子的平均智商以下。实际情况也确实如此。

9. 有一项关于种族血统的研究，对黑人孩子、白人孩子和黑人白人混血孩子进行了考察。这些孩子大都被中上阶层的白人家庭收养，黑人孩子的智商与黑人的整体智力水平没有太大区别，而黑人白人混血孩子的智商介于黑人孩子和白人孩子之间（Scarr and Weinberg，1976年和1983年）。

让我们按照顺序逐一对这些论点进行剖析。

智力的遗传度

《钟形曲线》一书中的论点似乎轻松而从容地结束了有关种族和智力的辩论。那些认为黑人与白人基因相同，或至少智力基因相同的人，一定相信黑人与白人没有什么不同，只不过他们所处的认知环境比较恶劣而已。如果实际情况确实如此，我们就需要测量环境对智力的影响，也就是说，我们需要找到两者之间的关联度。

双胞胎研究给出了智力和环境之间的关联度。赫恩斯坦和莫瑞假设了一个数值，他们认为这个数值代表的是环境对智力的最小影响，即他们假设60%的智力差异都是个体之间的基因差异造成的，剩下的40%是环境因素造成的。为了求得环境与智力之间的关联度，我们需要对环境造成的智力差异的百分比进行开平方运算。0.4的平方根约为0.63，这就是我们要寻找的智力和环境之间的关联度。现在我们就可以计算当黑人与白人的智力差距达到1个标准差时，普通黑人和普通白人所处的环境究竟相差多少。1个标准差除以0.63得到1.59个标准差，这就是这两个种族在环境方面存在的差异。简而言之，只有黑人所处的不利环境如此恶劣，才能够解释黑人与白人在智力上存在的差异！我们来看一下位于正态曲线的平均值下方的一组百分比，这样就能够理解这种环境差距究竟有多大。如果你找到低于平均值1.59个标准差的位置，你就会发现剩下的人口只有6%——这就是说，美国黑人所处的环境非常恶劣，只有6%的美国白人所处的环境在黑人的平均水平之下。

对于赫恩斯坦和莫瑞来说，这是解释不通的，但糟糕的还在后面呢。詹森

对双胞胎研究进行了最新的分析，这项分析显示，成年人的智力差异只有25%是环境因素造成的。0.25的平方根是0.5，1个标准差除以0.5得到2个标准差。现在我们就要假设黑人所处的平均环境比白人的差2个标准差，这就意味着只有2.27%的美国白人所处的环境在黑人的平均水平之下。而这是绝对不可能的。

这一论点的缺陷就在于，它假设同一群体内的环境因素和不同群体之间的环境因素是相同的。遗传学家理查德·莱旺庭曾经对这一缺陷进行过说明。想象将一袋种子随机分成A、B两组，这就意味着虽然同一组的种子之间也存在一定的基因差异，但这组种子之间平均来说不存在基因差异。将所有A组和B组的种子种在理想的培养土中，但是研究人员从B组中每一棵植株的培养土中去掉了一种营养成分。显然，在每一组内部，植株成熟时的高度完全由基因决定——毕竟在同一组中不存在环境差异。但是A、B两组植株高度的平均差异，就完全是由环境因素造成的，因为B组的培养土中缺少一种营养成分。

所以，我们现在似乎是证明了同一群体内的双胞胎研究（这些双胞胎中没有一个黑人一个白人的情况）并不能显示不同群体之间的环境因素带来的影响。但是还有一个问题，我们能用缺少一种营养成分的B组的环境和真实世界发生的情况进行类比吗？这种不利条件对B组的每个成员的影响必须同样大，但要对A组的每个成员都没有任何影响。即使是种族主义也无法达到这样绝对的效果。有些黑人比其他黑人在经济条件和其他方面的状况要好一些，有的白人也肯定在这些方面不如其他白人。莱旺庭的这种营养成分被命名为“X因素”，为的是最大程度地渲染这一因素的神秘效果，同时也表示没人能找到与其发挥同样效果的因素。

迪肯斯和弗林（2001年）提出了一个形式模型，解决了这一问题。他们证明了同一个具有极大影响力的环境因素可以将一群人分为两个群体，但是这个环境因素对不同群体内的每个成员的影响都是不一样的。他们用电视转播提升打篮球能力的例子来说明这一点。电视转播篮球赛使得篮球运动越来越流行，这就意味着年轻人打篮球的时间肯定更多一些，并且练就了一些新技术，例如过人或左右手都可以投篮等等。当毕业生重返学校和原来的高中篮球队进行比赛时，虽然他们在基因上不存在任何劣势（身高相仿，动作也一样迅速），但还是被打败了。显然，新的篮球环境并没有整齐地把这两个组分开，因为有些人在电视转播篮球赛之前，就已经开始刻苦练习，精进球技了。每组的不同成员受到的影响也是大小不均的。而在同一个年龄组内，和那些较矮、较壮的人相比，那些因遗传基因而更高、速度更快的人，更有可能喜欢篮球这项运动，并且能享受到团队合作的好处，得到教练更多的指导。这的确就是双胞胎研究让人模糊不清的地方。同卵双胞胎基因相同，身高和速度都相同，所以他们与篮球有关的经历就

很可能极其相似——而且既然基因和强大的环境因素造成了相同的结果，所有的原因就都被归结于基因的作用了。

现在我们很容易联想到一系列导致美国黑人与白人智力差距的环境因素。不同的养育方式、年轻人中流行的不同文化等，对每个群体都产生了强大的作用，影响了他们进行“智力练习”的数量以及他们发展解决问题的认知技能的程度。而且所有这些因素都不是像解释不通的X因素那样发挥作用。

在第6章中，我详细地介绍了黑人与白人之间存在的巨大环境差异。理论上是这些环境差异造成了不同种族在智力上的差异，现在谁也不能否认这一点。黑人与白人的智力差距由原来的1个标准差缩小至现在的2/3个标准差，这使我们更有理由相信，是环境造成了智力上的差异。当然，这种差距的缩小与基因没有任何关系。

与文化有关的测试题目和与文化无关的测试题目

黑人在宣称与文化无关的智力测试题目上的表现，不如在与文化有关的智力测试题目上的表现，从表面上看，这一论点似乎很难被驳倒。但是我们来回忆一下第3章（关于智力测试的分数随着时间的推移不断提高）中詹姆斯·弗林（2000年）的研究。这项研究显示，随着时间的推移，正是这些大家都认为“与文化无关”的智力测试和小测验的分数有了最大的提高。例如，被认为“与文化无关”的瑞文推理测验的分数一直以来不断上升，增长的幅度远远超过了明显与文化有更大关系的测试的分数。像词汇和信息这样的小测验分数，增长幅度远远不及积木图案（就是简单地对一些几何图形进行处理）这样宣称与文化无关的测试，或者其他测试操作智力或液态智力的小测验，例如物体拼组和图片排列。既然我们知道，基因不可能使智力在短短一代人的时间里实现大幅增长，那么我们就不得不认为是一些环境因素使得与文化无关的测验的分数有所增加，并且增加的幅度大于明显与文化有关的测验。所以，事实有力地反驳了智力与文化无关的论点。黑人与白人的得分相差最多的测试，恰恰是我们现在知道与文化关系最大的测试，例如瑞文推理测验和积木图案小测验。

撒哈拉沙漠以南的黑人的智商

现在我们暂停一下，花点儿时间想想，如果撒哈拉沙漠以南的黑人的实际智商真的只有70分，那么这么低的智商意味着什么呢？这一智商甚至低于智商最低的2%的白人。考虑一下，在我们的社会中智商这么低的人能够做些什么？我们会发现，普通非洲人或许根本不知道应该何时播种或者一个部落首领的职责是什么，也搞不清家庭成员的辈分关系。显然，这些非

洲人的智商一定是出了什么问题。这些非洲人的智商得分所代表的意义，一定不同于受欧洲文化影响的人的智商所代表的意义。

这么低的得分主要是从理查德·林恩和塔图·温汉南2000年总结的数据中得出的，参考的主要是（与环境高度呼应）的瑞文推理测验。林恩和温汉南使用的样本一般数量都很小，而且杂乱无章。不仅如此，他们还忽略了智商平均值较高的样本。所以这些测试的得分根本不能真实反映非洲人的智力水平。相反，这只能告诉我们，非洲人智力还没有取得大幅提高，尤其是瑞文推理测验所测试的液态智力。与此呼应的是，第3章中提到，最近的一项研究显示，在肯尼亚的某个地区，瑞文推理测验的分数在大约14年的时间里史无前例地提高了1.75个标准差。另外请注意，仅接受了几个月西式学校教育的非洲人，在液态智力测试中的分数就提高了0.5～0.7个标准差。即使是短时间的训练也能提高非洲黑人在瑞文推理测验中的分数，这些分数的提高相当于智商提高了14分（而这样的训练只将白人的分数提高了4分）。

赫恩斯坦和莫瑞（1994年）以及拉什顿和詹森（2005年）确认，美国“黑人”的基因中含有20%的欧洲基因。实际上，他们坚持认为，正因为这样，非裔美国人的平均智商才达到了85分，而非70分。当然，按照这个逻辑推理下去，如果黑人基因中含有的欧洲基因比例达到了40%而非20%，那么非裔美国人的平均智商就会达到100分——等同于白人的智商——而白人拥有的可是100%的欧洲基因。继续推理下去就会得出非常荒谬的结论，如果黑人基因中含有的欧洲基因达到了60%，那么非裔美国人的平均智商就会达到115分！

黑人在高智商水平试题上的表现更差

我们可以通过因素分析找出矩阵中各种相关因素保持一致的方式。当我们对智力测试中的题目（或小测验）进行因素分析时，第一个分离出来的因素叫作g因素。所有智力测试的题目（或小测验）都与这个因素相关。关于g因素一直存在着一个争议，那就是g因素的存在是否只是一种统计上的需要，即用g因素表示不同智力测试题目之间具有一定相关性。有些人认为g因素没有任何意义。有些人则将神经系统的生理性质作为依据，认为g因素具有重大意义。他们认为，g因素与一些生理或遗传变量有关，例如神经传导速度（NCV），这就证明g因素才是智力的主要源泉。赫恩斯坦和莫瑞（1994年）以及拉什顿和詹森（2005年）都认为，黑人与白人在具备较高g因素负荷（即与g因素的相关度）的智力测试题目或小测验中的分数差距更大，这就能够证明黑人与白人的智力差距是由生物因素和遗传基因造成的。

对于这样的论点，我们首先应该注意的就是，这一论点的主要依据就是韦克斯勒儿童智力测验量表中各项小测验的g负荷量。这一智力测试中各项小测验的g负荷量并没有多大差别，只有译码小测验的g负荷量偏低。除此之外，其他小测验的g负荷量差不多都在0.6～0.7之间。既然这项小测验的g负荷量并没有什么差别，怎么能说g负荷量与黑人和白人的智力差距有关系呢？

弗林曾经指出，有关g负荷量的论点的主要基础，即韦克斯勒儿童智力测验量表，在很大程度上向测量晶态智力的小测验倾斜，即信息、词汇、理解、算术和类同。如果某一类小测验在一项智力测试中占据了主导地位，那么第一个分离出来的因素在这类测验中的负荷量就会很高。所以，如果我们在智力测试中大量使用测量晶态智力的小测验，那么我们在因素分析中分离出来的第一个因素，也就是g因素，在很大程度上代表的就会是晶态智力。

但是，弗林还指出，假设我们不考虑晶态g因素而考虑液态g因素。詹森（1998年）等专家认为，遗传对液态g因素的影响至少等于遗传对晶态g因素的影响。要求受试者用几何图形拼出图案来，或是将图片按照有逻辑有因果关系的顺序排列起来，这样的小测验测量的都是液态g因素。詹森和其他智力领域的专家认为，瑞文推理测验测试的几乎是纯粹的液态g因素。所以我们就用韦克斯勒儿童智力测验量表的各项小测验与瑞文推理测验之间的相关度，来代表韦克斯勒儿童智力测验量表各项小测验中液态g因素的含量。现在我们将每项小测验中液态g因素的含量与弗林报道的最近几十年里每项智力小测验分数的增长幅度对比起来。结果我们发现，随着时间的推移，液态g因素负荷量越高的小测验，得分增加得越多。这样一来，就出现了自相矛盾的情况：智力的提高毫无疑问完全是环境造成的，而得分提高最多的小测验偏偏是那些被认为受到了最多遗传基因影响的小测验！因此，虽然我们知道了黑人与白人的智力差距在g负荷量较高的小测验和g负荷量较低的小测验上大小不等，我们却并不能由此确定黑人与白人的智力差距有多少是由基因造成的，有多少是由环境造成的。如果我们将g因素定义为晶态g因素，那么黑人与白人在g负荷量较高的小测验上的智力差距更大，这就应该能够表明这种差距是遗传基因引起的。但是如果我们将g因素定义为液态g因素，我们就会发现，某项小测验的g负荷量越高，受到环境变化的影响就越大。更容易受到基因影响的小测验不可能是受环境影响最大的小测验，所以很明显，关于g负荷的论点是错误的。

最后，正如我在第6章中指出的那样，有的论点认为某些试题g负荷量的高低能够显示黑人与白人间智力差距的大小，那么由此推断，在过去的30年里黑人在高g负荷量的试题上的得分应该增长得最少。威廉姆·迪肯斯和詹

姆斯·弗林（2006年）设计了一个“gQ测试”——是用韦克斯勒儿童智力测验量表中各项小测验的g负荷量加权后得到的智商。赫恩斯坦和莫瑞，以及拉什顿和詹森一定会说，即使黑人在智力测试中的分数能够提高5分，他们在g负荷量较高的测试中的分数也不会有多大的提高。而实际上，和白人相比，黑人在詹森定义的g负荷量较高的试题上的分数提高了5.13分。

现在我们应该完全清楚了，用关于g负荷的论点来解释黑人与白人的智力差距是由基因造成的，只是用来转移注意力的做法。各小测验的g负荷量没有什么区别，某一项小测验的g负荷量也不能证明黑人与白人的智力差距究竟是由生物因素、遗传因素造成的，还是由环境因素造成的。而黑人在g负荷量高的测试中的得分，和在g负荷量低的测试中的得分几乎有同样的增长。

黑人在受到近交衰退影响较大的小测验中表现更差

我们可以采取与反驳有关g负荷的观点相同的方式来反驳这一论点。从表面上看，如果近交衰退严重地影响了一项小测验的分数，并且如果黑人与白人极有可能在这项小测验上拉开很大的差距，那么这或许看上去能够很好地证明，黑人与白人的智力差距主要是由生物因素和遗传因素造成的结果。但实际情况却是，就像液态g因素负荷量一样，小测验受到的近交衰退的影响程度，同样与过去几十年里智商的增长幅度是相关的（弗林，2000年）。实际上，两者的关联度越高，智商增加的幅度就越大，黑人与白人之间的智力差距也就越大。这样一来，我们又陷入了自相矛盾的窘境。如果我们相信近交衰退的程度能够说明黑人与白人的智力差距是基因造成的，那么我们也就必须相信近交衰退的程度也能够说明智力的提高是基因造成的。换句话说，如果我们认可近交衰退的程度表明黑人与白人的智力差距是由生物因素造成的，那我们就必须相信近交衰退的程度也能够证明弗林效应——几代人的时间里智商逐渐增加——也是由生物因素造成的。

脑容积和黑人与白人的智力差异

白人的脑容积与智力之间的关联度大概是0.3～0.4之间（McDaniel，2005年；Schoenemann，Budinger，Sarich，Wang，1999年）。拉什顿和詹森（2005年）认为，黑人的平均脑容积小于白人。

不过，也不是所有的研究都发现黑人与白人的脑容积存在差别（美国国家航空航天局，1978年）。更重要的是，在白人中发现的脑容积和智力之间的相关性，并不意味着脑容积越大智商就越高。在同一个家庭里，平均来

说，脑容积较大的子女并不比其他脑容积较小的兄弟姐妹的智商高（Schoenemann，Budinger，Sarich，Wang，1999年）。

不管怎么说，正如我们前面曾经说过的，同一人口群中存在的差异并不一定能说明不同人口群之间存在这种差异的原因。在某一人口群中，更聪明的人脑容积更大，但并不能因此就认为，黑人与白人在脑容积上存在差异的原因，和同一种族内不同智力水平的人在脑容积上存在差异的原因相同。男性与女性在脑容积上存在的差异大于黑人与白人在脑容积上存在的差异（Ankney，1992年），但是女性与男性的平均智商是相同的。（我们应该注意到，通常智力测试会被设计成能够保证女性与男性的平均分数都是100分。不过女性与男性在大多数测试题目上的平均分数都相差无几，所以智力测试的设计能够保证性别平等并不是什么难事。）不仅如此，在厄瓜多尔有一群身材非常矮小的人，他们头部的大小比平均值低好几个标准差（Guevara-Aguire et al.，1991年）。这些人不仅个个智力正常，而且远远高于普通人（一多半的人在班级中都是最顶尖的学生）。

我们对大量黑人进行抽样后发现，黑人女性的脑容积与白人的脑容积一样大，但是两者的智力差距还是1个标准差，这也是收集数据时得出的典型的黑人与白人的智力差距。因此，即使是在脑容积相等的情况下，黑人与白人的智力仍然存在差距。

最后，虽然黑人与白人的脑容积大小不同，但有时候这种差异很可能是环境因素造成的，而不是基因引起的（Ho，Roessmann，Hause，Monroe，1981年）。和白人女性相比，黑人女性在怀孕期间更有可能遭遇到可能导致新生儿出生时体重过轻或脑容积较小的情况，例如营养不足和饮酒等（Bakalar，2007年）；而早产就与新生儿脑容积较小有关，尤其是黑人新生儿（Ho et al.，1981年）。只有当黑人新生儿早产时，他们的脑容积才会比白人新生儿小（Ho et al.，1981年）。黑人的产后条件也不如白人，尤其是营养条件（Ho，Roessmann，Straunfjord，Monroe，1980年）。

所以，尽管我们有时候发现黑人的脑容积小于白人，但我们从这一点上并不能得出什么令人信服的结论。我们不能从同一人口群中存在的关联推断不同人口群之间存在的差异，而且同一人口群之间的差异也许并不是由脑容积与高智商之间的因果关系而引起的。

黑人的反应时间更长

在白人中，高智商的人的反应时间比智商较低的人短。不仅如此，高智商的人的反应时间差别不大，也就是说，和智商较低的人相比，高智商的人的反应时间几乎是相同的。反应时间与智力的关联度很低——大约是

0.2（Deary，2001年），并且也不是总能发现两者之间存在联系，不过还是相信两者之间存在微弱的联系比较好。和白人相比，黑人的反应时间更长，并且不同黑人个体间的反应时间差异更大（拉什顿和詹森，2005年）。

同样，首先我们需要注意的是，在同一人口群中，对于脑容积或其他任何与智力有关的变量来说，不同人口群之间的差异，并不一定与同一人口群中存在的关联具有相同的原因。不仅如此，反应时间在11岁以后就不会再继续增加，而智商能一直迅速增长（Nettelbeck，1998年）。而且有些智力发育迟缓的人反应时间却极短（弗林，2007年）。

对于因反应时间长而认为黑人智商低的论点，这些还不是最大的问题。首先，赫恩斯坦和莫瑞（1994年），以及拉什顿和詹森（2005年）都坚持认为亚洲人的智商略高于白人，而且他们都暗示这种差距至少有一部分是基因造成的。在拉什顿和詹森从林恩和温汉南的一本著作中（2002年）援引的表I中，我们知道一个样本中的中国香港人平均智商达到113分，另一个样本中的日本人平均智商是110分（这样的智商比大部分文献中中国香港人和日本人的智商还高。顺便提一句，回忆一下第8章，有证据显示东亚人的智商并不比西方人高）。和表中的其他黑人与白人相比，东亚人的反应时间最短，并且反应时间的长短差别最小。但是詹森（1993年）称，和欧裔美国人相比，华裔美国人一组的反应时间更长，并且不同人的反应时间差异很大，尽管这群华裔美国人的智商比欧裔美国人高出5分。林恩和施戈西亚（1991年）报道称，虽然一组日本人的反应时间比一组英国人短，但不同日本人的反应时间差别比这组英国人要大。弗林（1991年）报道称，对于研究样本中的中国人来说，与智力相关的是动作时间，而不是反应时间。动作时间衡量的是一个人在决定移动手指以后将手指从起始位置移开所用的时间。大量研究都表明，动作时间与反应时间一样，都和智力高度相关（Deary，2001年），并且黑人的动作时间比白人短！但拉什顿和詹森（2005年）在他们讲述的有关反应时间与种族的简单故事里，对此只字未提。另外，根据林恩和温汉南的报道，南非黑人与爱尔兰人的智力差距很大，但两者的反应时间几乎没有什么差别。总之，所有的研究结果都与下面的论点不一致：①平均反应时间和反应时间的差异都与智力相关；②动作时间与智力不相关；③亚洲人的反应时间，而非动作时间，比白人短；④白人的反应时间比黑人短，但动作时间却比黑人长。这样一来，我们的结论就是，我们并不清楚反应时间、动作时间和种族之间的相互关系。

黑人的平均智商天生就比白人低？

遗传决定论者常常认为，由于基因的缘故，黑人的平均智商比白人低，如

果黑人孩子的父母与白人孩子的父母智商相同，那么黑人孩子的平均智商会降到白人孩子的平均智商以下。换句话说，黑人的高智商与黑人平均基因型之间的差距较大，而白人的高智商与白人平均基因型之间的差距较小，所以和父母为高智商白人的孩子相比，父母为高智商黑人的孩子的平均智商有着更大的下降空间。不仅如此，当黑人孩子的父母与白人孩子的父母智商相同时，黑人孩子的平均智商也低于白人孩子的平均智商。这种论点根本站不住脚，因为从环境决定论中我们可以得出相同的推论。如果环境因素，例如父母的养育行为和亚文化中不鼓励学业成就的压力，使得黑人的平均智商与白人的平均智商之间的差距拉得更大，那么黑人孩子的平均智商同样会降低到白人孩子的平均智商以下——而这与基因没有任何关系。

种族血统与智力差异

上面谈到的所有研究得出的结果都显示，黑人与白人的智力差距绝对不是由遗传基因造成的。但是这些证据多多少少都不具备足够的说服力，因为这些证据都是间接证据。在遗传这个问题上，唯一一个直接证据就是一个人的种族血统。美国黑人的基因中约有20%来自欧洲（Parra et al.，1998年；Parra，Kittles，Shriver，2004年）。一部分黑人的血统完全是非洲血统，不少黑人的基因中至少有一部分来自欧洲，还有一部分黑人——大约10%——的基因中占大多数的是欧洲基因。黑人基因中非洲基因和欧洲基因所占的比例大小，会使他们的智力有什么差别吗？如果按照遗传决定论的观点，基因中欧洲基因较多的黑人智商就较高。不过赫恩斯坦和莫瑞（1994年）以及拉什顿和詹森（2005年）恰恰几乎从来没有讨论过这一直接证据。

有关被白人家庭收养的不同种族血统的儿童的研究。赫恩斯坦和莫瑞（1994年），以及拉什顿和詹森（2005年）介绍了斯嘉和韦恩伯格（1983年）的一项研究。这项研究显示，被白人家庭收养的黑人儿童的平均智商，低于白人家庭收养的白人儿童的平均智商，被白人家庭收养的黑白混血儿童的平均智商则介于两者之间。根据黑人与白人的智力差距完全由基因所致的观点构建的最简单最典型的模型，被收养的白人儿童的平均智商应该比被收养的黑人儿童的平均智商高出15分或者更多，被白人家庭收养的黑白混血儿童的平均智商应该介于两者之间。当孩子大约7岁时，他们的智力几乎显示不出任何遗传的作用。但当他们成年后，他们的智力则受到了更多遗传因素的影响（韦恩伯格，斯嘉，沃德曼，1992年）。

斯嘉和韦恩伯格（1983年）认为，他们的这项研究中存在几个问题，使得这项研究不足以证明遗传假说。第一，收养机构有可能是有选择性地安置被收养人，因此黑人孩子有可能被安置在社会阶层较低的收养家庭里。第

二，既然我们并不知道被收养儿童的亲生父母的智商究竟是高是低，那么有可能被收养的白人孩子的亲生父母（的基因型）智商高于白人智商的整体水平，或者被收养的黑人孩子的亲生父母（的基因型）智商低于黑人智商的整体水平。这本身就可以说明为什么被收养的白人孩子平均智商高于被收养的黑人孩子的平均智商。第三，黑人孩子被收养时的年龄比白人孩子大，年龄偏大时被收养会对智力产生不利的影响。第四，与白人孩子相比，黑人孩子在进入收养家庭之前往往在数个不同的福利机构生活过，这也与较低的智力有关。第五，对黑人孩子在被收养之前的预适应安置情况也更差。第六，桑德拉·斯嘉告诉我，这些被收养的黑人孩子和黑白混血孩子，在青春期时会面临严重的心理障碍，而这些心理障碍都与身份问题有关。实际上，有一些孩子说："我朝镜子里望去，里面竟然站着一个黑人，我吃了一惊，因为我知道我真的是一个白人。"其他孩子也面临一些心理障碍，因为他们感觉他们真的是黑人，而为什么收养机构要将他们安置在一个陌生的白人收养家庭里？鉴于上述这些问题，斯嘉和韦恩伯格谨慎地劝告我们，不要从这项研究中得出任何有关遗传基因决定智力的结论。总之，正如我们现在所见，针对基因中含有欧洲基因能够使一个黑人更聪明这种论点，唯一一个对这一观点进行证明的研究就是斯嘉和韦恩伯格所做的混种族研究。

有关被收养的黑人儿童与白人儿童在有利于智力发展的环境中成长的研究。这是另外一项有关黑人儿童与白人儿童在同一环境中成长的研究。这项研究得出的结论与赫恩斯坦和莫瑞（1994年）以及拉什顿和詹森（2005年）根据斯嘉和韦恩伯格的研究所得出的结论截然不同。这项研究中的黑人儿童、白人儿童以及黑人与白人的混血儿童在某个环境优越的机构中成长。对孩子们进行照顾的人员都训练有素，并且态度非常认真。研究人员每天给孩子们安排的活动都是高度启发智力的活动。在四五岁时，白人儿童的平均智商是103分，黑人儿童的平均智商是108分，混血儿童的平均智商是106分。从表面上看来，这些结果似乎显示这些黑人儿童具有很明显的智力遗传优势。这项研究中的黑人儿童是西印度群岛人，白人儿童是英国人。虽然有可能是黑人儿童的父母拥有非常高的基因型智商，但弗林（1980年）指出，西印度群岛人有选择性地移民到英国最多只会使他们的智商分数提高有限的几分。尽管如此，与斯嘉和韦恩伯格的研究一样，我们依旧不知道被收养儿童的亲生父母的智商究竟是高是低。

有关被黑人或白人家庭收养的黑人儿童的研究。这项收养研究的设计与斯嘉和韦恩伯格的研究不同，但看起来显然更加合理。研究中的黑人孩子和黑白混血孩子，被中等阶层的黑人家庭或白人家庭收养（穆尔，1986年）。结果显示，无论是被黑人家庭还是白人家庭收养，这些孩子的智力都极为接近。因此，在黑人家庭或是白人家庭里，基因中含有欧洲基因并没有给被收养儿童带来什么优势。拉什顿和詹森（2005年）则认为这项研

究的结果并不成立，因为这项研究中孩子们在接受测试时才7岁。他们说："随着人们年龄的增加，基因会发挥越来越大的作用，而家庭的影响就会逐渐减弱。特征的差异在年岁较小时并不明显，但随后就会逐渐开始显现，到了17岁则完全表现出来。"但是他们研究的结果显示，孩子们7岁时和17岁时的遗传度完全相同，所以他们自己的证据都证明了穆尔的研究发现是成立的，即黑人儿童与混血儿童的智力没有差异。不仅如此，更宽泛地说，还有很多证据显示儿童到7岁时的智力遗传度就已经很高了，所以在这个年龄上，不同种族的智力没有差异是很能说明问题的。

穆尔的研究为我们证明环境假说和遗传假说奠定了基础。如果假设种族之间的智力差距主要是由基因造成的，那么被收养的黑人儿童和混血儿童，无论是在黑人家庭还是在白人家庭内长大，他们的智力都不会有什么差别。如果假设种族之间的智力差距主要是由环境造成的，例如家庭、社区和学校的环境，那么这些孩子是在黑人家庭还是在白人家庭内长大，他们的智力就应该有很大的差别。因此，即使黑人收养家庭和白人收养家庭都属于中等阶层，研究人员仍然预测被白人家庭收养的孩子的智商更高。实际情况也确实如此。被黑人家庭收养的孩子平均智商是104分，而被白人家庭收养的孩子平均智商是117分。总之，这项研究并没有证明黑人与白人的智力差距是遗传基因造成的，而是让我们有充足的理由相信这样的差距主要是或者完全是环境因素造成的。不过，在这项研究中我们并不知道被收养儿童的亲生父母的智商究竟是多少（并且样本的数量——46人——也不太令人满意）。

幸好还有其他一些研究能够让我们比收养研究更直接地评估欧洲血统和非洲血统的影响。美国所有黑人的基因都是介于100%的西非基因到大部分都是欧洲基因之间。那么拥有欧洲基因的黑人智商会更高吗？对这个问题，有5项不同的定性研究分别给出了回答。

有关肤色的研究。研究肤色深浅和智力高低的研究很容易进行，所以以往有很多这样的研究。让我们先暂停一下，思考这样一个问题：假设黑人与白人的智力差异完全是环境造成的，那么在黑人的肤色和智力之间应该具备何种关联度呢？我们应该发现，肤色较浅的非裔美国人智商更高，并且因此取得更高的社会经济地位以及教育和环境方面的优势。这样的话，肤色和智力之间的关联度应该在0.2~0.3甚至更高。而实际上所有关于这类研究的介绍都一致表明，黑人的肤色与智力之间的关联度很低。即使是奥德丽·舒伊（Audrey Shuey）——最坚信黑人与白人的智力差异是基因所致的人，得出的结论也是一样的。黑人的肤色与智力之间典型的相关度是0.1~0.15。五官与典型非洲人五官之间的相似程度，与智力之间的关系同样很低（Shuey，1966年）。即使将肤色较浅有可能给黑人带来的优势忽略不计，仅为0.1的关联度也不能说明欧洲血统会对一个黑人的智力发

挥强大的遗传作用。另一方面，舒伊回顾的很多研究样本数量都很少，并且选取样本的方式也不能令人信服。虽然这些研究对肤色和智力的测量都很准确，但主要的问题是，看上去肤色似乎可以直观地反映一个黑人拥有多少欧洲血统，但实际上并非如此。撒哈拉沙漠以南的黑人，肤色的深浅差异很大，所以一些非洲人虽然没有欧洲血统但肤色却较浅。因此，如果要证明关于“欧洲血统”的假说，就需要更加可靠的指标来衡量一个黑人是否具有欧洲血统。

通过与血型有关的指标衡量黑人是否具有欧洲血统的研究。有关肤色与智力的研究没能得出任何结论，但幸好还有其他数据能够继续对有关欧洲血统的假说加以证明。不同血型在不同种族中的比例各不相同。有些在欧洲人口中很常见的血型在非洲人口中就很稀有，反之亦然。在遗传决定智力的假设下，具有更多“欧洲”血型的黑人应该拥有更多的欧洲基因，因此智商也就更高。但是斯嘉、帕克斯蒂斯、卡茨和巴克（1977年）发现，在一个由144名黑人双胞胎青少年构成的样本中，智力与黑人拥有欧洲遗传基因的多少之间的关联度只有0.05。如果不考虑肤色和社会经济地位，这一关联度还会降至0.02。值得注意的是，这些研究人员发现肤色与智力之间的典型相关度是0.15。这就意味着，如果一些研究得出的相关度也在0.15左右，那么这一相关度代表的并不是欧洲基因与智力的关系，而是其他一些与浅肤色有关的因素与智力的关系，例如浅肤色给黑人带来的社会优势等。

洛林及其同事（1973年）选取了由少量黑人构成的两个不同的样本，然后寻找估测到的血型的欧洲特征（而不是个人的欧洲特征。通过对血型的估测得出一个人具备欧洲特征的多少）与智力之间的相关度。一个样本得出的相关度是0.01，而另一个样本得出了一个没有意义的相关度——0.38，也就是说，与非洲人的血型更接近的黑人智商更高。

尽管如此，我们应该注意到，关于血型的研究并不像表面上看起来那样确定。这是一些技术原因造成的。这些技术原因与一个情况有关，即在黑人中，白人的血型基因间的联系极其微弱。如果白人的血型基因彼此之间不存在联系，那么或许它们与决定智力高低的白人基因也没有联系。

有关第二次世界大战期间美国黑人士兵与白人士兵的子女的研究。一位德国心理学家（Eyferth，1961年）对数百个孩子的智力进行了研究，这些孩子都是1945年占领德国以后美国黑人士兵与当地的德国女性生育的私生子。将这些孩子的智力与当时美国白人士兵与当地德国女性生育的私生子的智力进行比较。我们可以先在大脑里想象一下这个实验。我们知道，这些美国黑人与德国女性所生的混血孩子会遭遇严重的偏见，因为显而易见他们是德国女性与外国士兵生育的私生子。这样一来，即使我们假设黑人与白人的智力差异与基因没有任何关系，这些混血孩子所处的不利

地位也决定了他们的智商不会很高。但实际上，美国黑人士兵的私生子女的平均智商是96.5分，美国白人士兵的私生子女的平均智力是97分。由于军队里黑人与白人的智力差距很接近黑人与白人的整体智力差距，所以这些数据说明，美国人口中黑人与白人的整体智力差距并不是遗传造成的。不过这些数据也不像表面上看起来那样有说服力，因为美国军队在录用士兵时对智商有最低要求。这一要求将更多的黑人拒之门外，这也就意味着录用的黑人士兵并不能代表黑人的整体智力水平。虽然据弗林估计，美国军队中黑人士兵的智商只比黑人的整体智商水平高出3分或者更少，但这一漏洞表明，这项研究的结果并不那么确定。值得注意的是，有些私生子的父亲是北非军队的士兵。尽管如此，据弗林（1980年）估计，这几乎不会影响我们预测的黑人士兵的私生子的智商——除非我们假设北非士兵的平均基因型智商比已知的任何军队的士兵都高。

白人血统的作用。第三个查明黑人拥有的白人血统的方法就是询问他们的家族历史。假设黑人与白人智商相差的15分主要是由基因造成的。然后我们来考虑四类黑人：只有非洲血统的黑人；非洲血统多于白人血统的黑人；非洲血统和白人血统一样多的黑人；白人血统多于非洲血统的黑人。在黑人与白人的智力差距由基因所致的假设前提下，这四类黑人的智商也应该不尽相同。如果我们将智商非常高的黑人单独挑出来，我们应该发现其中有很多人都拥有很大比例的白人血统，而且这些人所占的比例远远大于拥有较大比例白人血统的黑人在整个黑人人口中的比例。

威蒂和詹金斯（1934年和1936年）从芝加哥一所学校中的黑人小学生中挑选了一个研究样本，其中有63名学生的智商是125分或125分以上，28名学生的智商在140分或以上。通过他们对自己家族血统的介绍，调查人员将他们按照上面的四种血统类型分类。智商是125分或以上的学生和智商在140分或以上的学生，拥有的欧洲血统都略少于对美国黑人整体水平的最佳估计值。不过，这项研究也存在一定的问题。最好的办法应该是将芝加哥地区高智商的黑人小学生所拥有的欧洲血统的多少，和这一地区其他黑人小学生进行比较，而不是和黑人人口的整体水平进行比较。不过这项研究的结果再次说明了，黑人与白人的智力差距和遗传基因没有任何关系，或许还能说明非洲人具备微小的遗传优势。

比较白人母亲和黑人母亲生育的混血孩子的研究。如果黑人与白人的智力差距主要是由遗传基因造成的，那么不管混血孩子的父亲是黑人，还是母亲是黑人，他们的平均智商应该是相同的，因为不存在先验理由让我们认为母亲是黑人、父亲是白人的孩子的基因型不同于母亲是白人、父亲是黑人的孩子的基因型。（不过需要注意的是，研究中的黑人父亲受教育程度和工作的级别都高于一般的黑人男性。这就是说，母亲是白人、父亲是黑人的孩子的基因型略好于母亲是黑人、父亲是白人的孩子的基因型。）

但是，如果：①在对孩子进行智力启发的过程中，母亲的作用大于父亲的作用，并且如果白人父母的养育方式鼓励孩子们获得了各种有利于提高智力的技能；②如果因为孩子的母亲是白人，孩子接触的同龄人更多的是白人。那么母亲是白人、父亲是黑人的孩子的智商应该高于母亲是黑人、父亲是白人的孩子的智商。实际上，威勒曼及其同事（1974年）进行的一项研究发现，母亲是白人、父亲是黑人的孩子的智商，比母亲是黑人、父亲是白人的孩子的智商高9分。这一结果说明，黑人与白人的智力差距绝大部分是由环境因素所致（不过请注意，接受测试的孩子只有4岁，根据这个年纪测得的智商预测成年后的智商，准确度非常有限）。

那么在有关种族血统的研究——到目前为止，这是评估黑人与白人的智力差距究竟是遗传所致还是环境所致的最直接的方法——我们从中都得出了哪些结论呢？一项有缺陷的收养研究显示，这一差距主要是基因所致；还有两个略有瑕疵的研究，一个显示非洲的基因比白人的基因略微优秀一点，另一个则显示两者的基因没有差别。还有很多根据肤色和黑人的五官特征来判断种族血统的研究，都不能为遗传决定论提供足够的支持。不仅如此，有三项关于拥有不同程度欧洲血统的研究，分别采用了两种不同的设计，也都不能证明遗传决定论的观点。关于德国私生子的研究并没有显示父亲为白人的孩子比父亲为黑人的孩子智商高。一项研究表明，智商极高的黑人孩子的血统中，欧洲血统所占的比例并不高于根据对黑人人口的整体水平做出的最佳估计值。还有一项研究显示，对于混血孩子来说，母亲为白人比母亲为黑人更有利于孩子的智力发展。

所有这些有关种族血统的研究都可以从不同的角度来解释。不过这些解释大都归结为一点，即黑人与白人的结合中，智力存在自我选择的可能性。如果与黑人结合的白人的智商远远低于一般白人的智商，那么他们的欧洲基因就不会使子女具备任何智力优势。同样，如果与白人结合的黑人的智商远远高于一般黑人的智商，那么他们的非洲基因也不会使子女具备任何智力劣势。然而在混种族的结合中，白人基因的贡献要小于一般的白人基因，或者黑人基因的贡献要大于一般的黑人基因，并且这种小或大的幅度必须在极端的情况下，才会使完全具备非洲血统的孩子和具有部分欧洲血统的孩子的智力完全相等。不仅如此，在大多数黑人与白人结合的奴隶制时代，智力的自我选择不太可能具有很大的效应。例如，与黑人女性结合的白人男性的平均智商不太可能低于其他白人男性的智商。的确，如果这样的结合大多是白人奴隶主与黑人女奴之间的结合，实际情况很有可能就是这样（Parra et al.，1988年），如果与智力相关的经济地位较为有利（就像现在的实际情况一样），这些白人的智商很可能比白人的平均智商略低。这类结合中的白人男性在选择黑人女性时的依据是她们容貌的美丑，而非智商的高低。同样，“二战”时的美国士兵，不论是黑人还是白人，在选择德国女性时的标准也不太可能是智商的高低。

另外，还有好几项研究完全排除了自我选择的效应。尤其是关于在某个机构中成长的黑人和白人儿童的研究，以及被中等阶层黑人或白人家庭收养的黑人儿童的研究，都不能用智力的自我选择效应来解释。

总之，虽然阅读赫恩斯坦和莫瑞的书（1994年），以及拉什顿和詹森的文章（2005年），我们永远无法得出有关种族血统和智力关系的确切结论，但这些纷繁复杂的证据——也是我们拥有的唯一的直接证据——全都无法证明黑人与白人的智力差异是遗传基因造成的。

黑人的智力提高了吗？

虽然我们有证据能够有力地证明黑人与白人的智力差距并非基因所致，但是如果没有证据显示在过去一代人的时间里智力有所增加的话，我们还是不能完全放弃遗传决定论的假说。随着民权运动、机会均等行动计划的兴起，更多的黑人迈入中等阶层的行列，黑人已经渗透到社会的最高等级（包括一位参谋长联席会议主席，两位国务卿，一位强有力的总统，全球最大的媒体公司的首席执行官，全球最大的经纪公司之一的首席执行官），黑人的物质条件和社会条件都得到了一定的改善。伴随着这些变化，黑人的平均智商有没有提高呢？拉什顿和詹森在2005年发表的文章中，给出的是否定的答案。他们认为，黑人与白人的智力差距在100年来都稳定地保持在1.1个标准差，或者16.5分左右。

但是威廉姆·迪肯斯和詹姆斯·弗林证明，通过对不足25岁的受试者进行智力测试，得出的结果显示，1972—2002年，美国黑人与非拉美裔美国白人的智商差距缩小了4.5~7分（使用的测试的种类不同，其差距也不同，迪肯斯和弗林更倾向于5.5分）。正如我在第3章中指出的那样，智力测试需要不断更新进行标准化，否则就会过时。测试人员需要尽可能地随机选取研究对象构成标准版本研究的样本。迪肯斯和弗林收集了四项使用最广泛的智力测试曾经使用过的9个标准版本：韦克斯勒儿童智力测验量表，韦克斯勒成人智力测验量表，斯坦福–比奈智力测验量表以及武装部队职业资格测试。他们认为这几项智力测试最能有效地估测黑人与白人的智力差距所发生的变化。收集到的韦克斯勒儿童智力测验量表过去整整30年的数据显示，黑人与白人的智商差距缩小了5.5分。其他三项智力测试标准化的时间间隔更短，不过平均来说，对这些数据的分析做出的预测也得出了相同的结论。在30年的时间里，黑人与白人的智力差距缩小了1/3个标准差，或者比30年前缩小了1/3。

迪肯斯和弗林没有采用其他5项智力测试的数据，虽然这几项测试也完全实现了更新，但他们认为这几项研究的样本或设计存在缺陷。拉什顿和詹森却认为应该将其中4项测试的数据包括在内，而所有这4项测试的数据显

示的黑人与白人智力差距缩小的幅度，都小于迪肯斯和弗林的研究结果。尽管如此，如果我们把所有这9项测试的数据都算在内，虽然有一部分测试的数据显示差距只减少了2～3分，但所有测试数据的平均值也有4.5分，这与迪肯斯和弗林的估计并没有太大差别。

如果全部人口的整体智商，在每30年一代人的时间里平均增加了9分，并且黑人与白人的智商差距在过去的30年里减少了5分，我们会发现，黑人现在的智商要高于白人以前的智商。弗林曾经提出这样一个问题，如果现在的黑人参加1947—1948年第一版的韦克斯勒成人智力测验量表（标准化时的样本中全部是白人）的测试，他们的得分会是多少？那是近60年前的事了，在这段时间里全部人口的平均智商增加了18分。弗林计算，现在黑人的得分要比1947—1948年的白人高出4分。

黑人与白人的智力差距完全是环境造成的

随着时间的推移，大幅缩小的不光是黑人与白人的智力差距。据非常可靠的数据显示，黑人与白人在阅读和数学能力上的差距也缩小了。阅读和数学能力都是反映智力高低的重要指标，和智商具有同样重要的地位。每隔几年，美国教育部就会随机选择一些9岁、13岁和17岁的孩子，参加一项叫作美国教育进展评估的测试。我们可以观察出生于1954—1994年的黑人孩子与白人在阅读和数学能力方面的差距。测试初期，黑人孩子的阅读成绩远远低于白人孩子，两者的平均差距是1.1～1.2个标准差。最近的测试则显示，两者的平均差距是0.6～0.8个标准差——差距缩小的幅度非常大。需要注意的是，在这段时间里黑人孩子取得的进步是不均匀的。早期差距以惊人的速度不断缩小，中期时差距又出现了不断扩大的趋势。只是从最近开始，差距才开始重新出现缩小的势头。为什么20世纪50年代中期到70年代初期出生的孩子进步的速度如此之快，或者为什么稍晚一些出生的孩子进步的速度发生了变化，或者为什么又重新出现了进步的趋势，对于这些，我没有什么可以让人信服的解释。

关于黑人与白人在数学成绩上的差距也有好消息。最早进行测试的孩子中，黑人与白人的差距足足达到了1.2个标准差。出生于20世纪50年代中期到60年代末期的孩子，在数学方面取得了惊人的进步，出生于20世纪70年代初期和80年代末期的孩子的差距保持不变或略微有所增加，此后出生的孩子的数学成绩差距再度缩小。同样，也没有明确的解释能够说明，为什么发展趋势会发生这样的变化，不过总体情况还是好的：数学成绩的差距在0.6～0.9之间，比原来缩小了1/3还要多。

有趣的是，如果我们将美国教育进展评估测试中取得的进步，按照智商量表进行换算，即平均值设定为100分，标准差设定为15分，并对9岁、13

岁和17岁儿童在阅读和数学方面取得的进步分别求平均值，我们就会发现，在迪肯斯和弗林（2006年）发现黑人与白人的智商差距缩小了5.5分的那个时间段里，黑人与白人在阅读和数学成绩上的差距缩小了5.4分。

总之，黑人与白人在智力上的差距由基因所致的论点，没有任何直接证据，这一论点本身也站不住脚，很容易就可以被驳倒。最直接的证据——也是唯一真正能够算做证据的证据——与黑人的欧洲血统有关。在这唯一一个例外中——赫恩斯坦和莫瑞（1994年）或拉什顿和詹森（2005年），恰巧都没有给出关于这项研究的详细报告——数据显示，更多的欧洲基因并不会给一个黑人带来任何智力上的优势。在过去的30年里，黑人与白人在智力上和学习成绩上的差距都缩小了大约1/3。这些证据都证明，黑人与白人之间既存的智力差异完全是环境造成的。

致谢

本书的写作以及部分书中提及的研究，都得到了美国国家科学基金会（National Science Foundation，授权号0717982）和美国国家老龄化研究所（National Institute on Aging，授权号1RO1AG029509-01A2）的大力支持。不过本书所表达的观点并不代表这两个机构的看法。哥伦比亚大学心理学系以及拉塞尔·塞奇基金会（Russell Sage Foundation）也为本书提供了种种宝贵的资源和设施。

很多人都慷慨地提供了他们的看法和意见，为本书增色不少——不过书中若是出现任何错误，他们都不用负责。这些人包括乔舒亚·阿伦森、道格拉斯·贝沙罗夫、克兰西·布莱尔、珍妮·布鲁克斯—冈恩、汉娜·查、威廉姆·迪肯斯、詹姆斯·R. 弗林、菲利浦·戈夫、理查德·冈萨雷斯、大卫·戈理斯莫、戴安娜·哈尔彭、劳伦斯·赫希菲尔德、厄尔·亨特、欣诺布·基达雅玛、马特·麦古、沃尔特·米歇尔、甘道夫·内斯、丹·奥舍森、达芙娜·奥伊瑟曼、丹尼斯·帕克、理查德·罗思坦、彼得·萨洛维、肯尼斯·萨维茨基、爱德华·E. 史密斯、雅基·史密斯、克劳德·斯蒂尔、罗伯特·斯坦伯格、艾瑞克·特克海默、芭芭拉·特韦尔斯基、简·沃尔德夫格尔以及奥斯卡·伊巴拉。我很感谢我的代理人约翰·布罗克曼和凯廷卡·马特森，是他们出面替我料理相关事务，使得这本科学读物能够顺利地与公众见面。我同样要感谢编辑安琪拉·冯·德·利普、埃丽卡·斯特恩，以及玛丽·巴布科克，她们出色地完成了编辑工作，使这本书得以出版。劳拉·雷诺兹帮助我整理了手稿。凯瑟琳·赖斯不仅帮忙在图书馆查阅了大量资料，还热心地为本书提供了不少建设性的意见。苏珊·尼斯贝特也给出了不少颇有见地的建议。

另外，李·罗斯也为本书做了很多贡献。我在读研究生时第一次遇到他，后来我所有的项目都得到了他提供的大量帮助。谨以此书感谢他对我的启发和关爱。